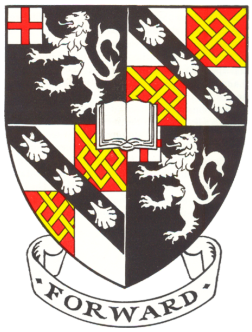
Optical fibre communication over a noisy partially coherent channel



Martín Sales Llopis

Department of Engineering
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Churchill College

October 2019

Declaration

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as specified in the text. It does not exceed the prescribed word limit for the relevant Degree Committee. This dissertation contains fewer than 50,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 80 figures.

Martín Sales Llopis

Abstract

As global IP traffic grows unceasingly, optical networks demand for technology upgrades in order to keep the feared "capacity crunch" away. The most celebrated technologies of coherent detection and wavelength-division multiplexing (WDM), widely deployed in long-haul links, are gaining ground in access networks, which is particularly challenging due to the shared-cost requirements, leading to denser channel spacings and the use of cheaper devices that tend to be noisier. In order to make the most of this technology combination, it is crucial to have a model of the channel that accurately describes all the present sources of noise. Traditionally, the most used model has been the additive white Gaussian noise (AWGN) channel, which, although only accounting for a linear contribution of complex noise and being insensitive to rotational phenomena, has shown its validity in numerous studies, as well as in commercial equipment. In this thesis, however, it is observed that the adoption of coherent detection and WDM, with lower-grade semiconductor lasers showing a moderate linewidth, yields scenarios where a phase-sensitive model becomes a must. The partially coherent AWGN (PCAWGN) channel is a popular choice that fulfils this need, but its high complexity due to non-trivial functions involved, deprives it from being suitable in high-speed digital circuits. The main goal of this thesis is to describe a reduced-complexity approximation in polar coordinates, accurate enough to find its applicability in modern systems. Furthermore, this works explores some possible end-to-end applications, like channel capacity estimation or symbol detection, assessing its performance by means of extensive simulations. Lastly, the emerging field of complex modulation of directly modulated lasers is revisited, with a special interest in how the proposed approximation can help to improve the performance of previously reported techniques, as well as proposing a new way to design spiral-shaped constellations aimed to maximise the channel capacity.

Title: "Optical fibre communication over a noisy partially coherent channel"

Author: Martín Sales Llopis

To my nephew, Jesús

Acknowledgements

First and foremost, I would like to express my most sincere recognition to my supervisor, Prof Seb J. Savory, whose infinite support, wisdom, and patience can only be rewarded with eternal gratitude. Thank you, Seb, for taking me under your wing, believing in my capacity to undertake research, and, in essence, for everything you have helped me with during the ups and downs of the last five years.

I am also very grateful to my advisor, Prof Richard Penty, for being very supportive since the beginning of my stay in Cambridge, and for giving me the opportunity to participate in supervision sessions that have fructified in a valuable teaching experience for me.

It has been a pleasure to work with the rest of the team at CAPE: David (thank you for the interesting discussions), Robert, Jhon, and my good friends Javier and Faruk. A big thank you also to Adrian Wonfor and Victoria Rose, for all the computing and administrative support, respectively, to Nikos for his help with lab demonstrations (and his football skills), and to Marcin (very especially), Ilana, and Eric for the most enjoyable times in the department (and sometimes out of it).

From my teething period at UCL, I thank Prof Polina Bayvel for the warm welcome and, particularly, a special word to Dr Domaniç Lavery (Dom), whom I learnt so many things from and whom I feel true admiration for. Also in UCL, I was lucky to share the working space with excellent people: malaka Paris and the great Rachid (I am so fond of you, guys), Zhe, Kari, Joshua, and the cool people from "the other room": Alex, Nikita, Gabriele, Gabriel, and Lidia. Then, there is Sezer, who has been a reference to me and has become an appreciated friend of mine. Thank you, Sezer, for the good times and for having my back whenever I needed.

However topical it might sound, nothing of this would have been possible without the initiative of the European Commission´s Marie Skłodowska-Curie Actions program, in my case within the framework of the ICONE project. Apart from the

financial support —for which I could not be more grateful—, it has been amazing to be part of such an spectacular team of people, amongst whom I want to highlight Prof Andrew Ellis' leadership, and Noreen's hard work at the organisation. Special thanks to Dr Paul Harper, for hosting me in Aston University and for his enthusiastic explanations. And, of course, the gang: Oskars, Tu, Peppe, Jaime, Aleks, Auro, Simone, Francesca, Hugo, Asif, Aditya, and Kseniia, who did a superb job as our student representative. What a great time we've had together!.

Although we met time ago in UCL (as he was part of the committee that first interviewed me), it was during my recent internship in Microsoft that I had the privilege of being tutored by Dr Benn Thomsen, whom I consider an example of intelligence, charisma, humbleness, and he is such a good listener, too. It was also during this time, that I really got to know Kai. I am so glad I had your advice, your valuable insight in our beer-enhanced discussions, and, of course, all the times we enjoyed playing football and tennis matches.

Being away from home for such a long time is not easy. Friends are the part of life that makes it liveable. So I started calling home wherever I found my friends. It was first in London, where I was blessed with the inestimable help of Carlos (Mousy), one of the best persons and friends that I have ever met. Then, home was moved to Cambridge, where Churchill College was undoubtedly the place to be, and where I treasured friends like Jack (rocket man), Charlie, Holly, Thiago, Fabian, Sasha, Janosz, Max, Nadi, Chiara, and, in general, all the good people I met in the MCR. Out of college, I found wonderful people too, like Unai, whom I owe so much for his help and generosity during my move out (eskerrik asko, Unai).

With the lapse of years, as one moves around, good friends are met and left, then re-encountered, then left again. But they are always there, happy to see you, no matter how long since the last time. Amongst these, spread around the world, I am fortunate to have friends like Rahul and Vaishali, Alicia, Tomeu, Julià, Shufen, Robert, Elene, Miguel, and Sabrina. All of them, in their own way, contributed to supporting me with this PhD.

I wish to dedicate these lines to Susannah, for being by my side along these unforgettable years together, for all that we shared, and no matter where life takes us from now. You really deserve a master degree in being patient and tolerant with me. I am also so very grateful to the whole Evans family, particularly to mum Liz and dad Richard, for making me feel like one more at home in Alty.

Finally, last but not least, I want to address my old friends and my beloved family, allowing myself the license to use the same languages that I would use to speak to them, Spanish and Majorcan.

Lo mejor de estos interminables años siempre ha sido el retorno. Desde luego, no hay nada como volver a "sa roqueta" y encontrarme con mis amigxs, lxs de siempre: Carlos (el hermano que nunca tuve, pero que sí tengo, porque él existe y siempre está ahí), Teru (que va muy torrado), David (pa la finca!), Simó (que siempre monea), Sureda (no, si se lo comerá), Pablo (un chinchón, padre), Joan (uep, bones!), Lluis (he tú Lluis, he tú), Lillo (polla de barro), Chevi (sabrosón, asúcar), Manolo (Manolario), y Jesús (ezúuu). Y, por supuesto, no me olvidaría jamás de mis estimadísimas Lu, Mar, Margui, Irene, Raquel, Juana, Pepa, Mira, Auba, Marta y Sofi. Gràcies guapos i guapes, per tot! — *Vos estim a tots igual* :)

Mi más sincero y profundo agradecimiento es, de todo corazón, para mi familia, única e irrepetible. Gracias Papá y Mamá, a mis hermanas, Mari Carmen y Paquita. También a mis primos Jaume y Tià, a *ses nines*, Joana i Leonor, a mi tío Agustí, a mi madrina Sebastiana y a mi cuñado, Jesús Miguel (guapísimo). Por el amor incondicional, por todo el apoyo durante tantos (y tantos) años, por haber creído siempre en mí. Siento mucho la ausencia, tan larga, los fines de semana tan esporádicos, así que gracias por las esperas, vuestras visitas y los continuos ánimos para seguir peleando por ello. ¡Os quiero mucho!.

I, per acabar, molt, molt especialment, seré eternament agraït de tenir a sa millor pradina del món sencer, sa matriarca, ningú cuida, ni cuina com tu. Segur que vares resar un parell de rosaris perquè jo pogués acabar i pareix que encara anirà bé. T'estim molt *païna*

Contents

List of Figures

List of Tables

Acronyms

ADC	Analogue-to-Digital Converter.
AIR	Achievable Information Rate.
APP	*A Posteriori* Probability.
ASE	Amplified Spontaneous Emission.
ASIC	Application-Specific Integrated Circuit.
AWGN	Additive White Gaussian Noise.
BER	Bit Error Rate.
BPS	Blind Phase Search.
BPSK	Binary Phase Shift Keying.
CD	Chromatic Dispersion.
CM	Complex Modulation.
CMA	Constant Modulus Algorithm.
CORDIC	COordinate Rotation DIgital Computer.
CPE	Carrier Phase Estimation.
CS	Constellation Shaping.
CW	Continuous Wave.
DAC	Digital-to-Analogue Converter.
DC	Direct Current.
DD	Decision Directed.
DFB	Distributed Feedback Laser.

DML Directly Modulated Laser.

DSP Digital Signal Processing.

DWDM Dense Wavelength Division Multiplexing.

EAM Electro-Absorption Modulator.

EDFA Erbium Doped Fibre Amplifier.

EML Externally Modulated Laser.

EON Elastic Optical Network.

FIR Finite Impulse Response.

FPGA Field-Programmable Gate Array.

FSK Frequency Shift Keying.

FTTx Fibre to the Home, Cabinet, Curb, Office, Node, Building.

GVD Group Velocity Dispersion.

HD Hard-Decision.

IMDD Intensity-modulated Direct Detection.

IQM In-Phase and Quadrature Modulator.

ISI Inter-Symbol Interference.

IT Information Theory.

ITU International Telecommunication Union.

LO Local Oscillator.

LUT Look-Up Table.

MAP Maximum *A Posteriori*.

MI Mutual Information.

MIMO Multiple Input Multiple Output.

MLSE Maximum-Likelihood Sequence Estimation.

MMF MultiMode Fibre.

MZM Mach-Zehnder Modulator.

NDA Non-Data Aided.

NGOA Next Generation Optical Access.

NLPN NonLinear Phase Noise.

OAN Optical Access Network.

OLT Optical Line Terminal.

ONU Optical Network Unit.

OOK On-Off Keying.

PAM Pulse Amplitude Modulation.

PBS Polarisation Beam Splitter.

PCAWGN Partially Coherent Additive White Gaussian Noise channel.

PDF Probability Density Function.

PLL Phase-Locked Loop.

PM Phase Modulator.

PMD Polarisation Mode Dispersion.

PON Passive Optical Network.

PSK Phase Shift Keying.

QAM Quadrature Amplitude Modulation.

QPSK Quadrature Phase-Shift Keyed.

RF Radio Frequency.

SAPM Simultaneous Amplitude and Phase Modulation.

SER Symbol Error Rate.

SMF Single-Mode Fibre.

SNR Signal-to-Noise Ratio.

SOP State of Polarisation.

SPSK Spiral Phase Shift Keying.

TDM Time Division Multiplexing.

TIA TransImpedance Amplifier.

uDWDM ultra-Dense Wavelength Division Multiplexing.

VA Viterbi Algorithm.

VOA Variable Optical Attenuator.

WDM Wavelength Division Multiplexing.

XPM Cross-Phase Modulation.

Chapter 1

Introduction

1.1 Facing the capacity crunch

Since the introduction of commercial optical systems in the 1980's, the capacity of fibre-optic systems has grown 10-fold every 3 to 5 years thanks to several achievements such as improved optical fibres, the erbium doped fibre amplifier (EDFA), wavelength-division multiplexing technology (WDM), and high-spectral efficiency coding. During the same period, the actual traffic demand grew by the same amount only every 7 years [1]. Notably, the telecommunications industry has comfortably provided technologies that satisfied the bandwidth demand for more than thirty years. Figure 1.1 illustrates the chronological evolution of such technologies along with the corresponding global traffic demand.

However, while both industry and research teams are witnessing how current technologies are approaching the nonlinear capacity limit of optical fibre [2], the global Internet generated traffic is expected to grow at a compound rate of 22% per year until 2022 [3]. This growth is mainly fuelled by the rapid development of the Internet of Things (IoT) and the proliferation of media and entertaining content platforms offering immediate streaming of high-resolution video, such as 4K or 8K. To give some forecast figures, by 2022 the global network is expected to host 4.8 billion users, 28.5 billion connected devices and 82% of the overall IP traffic will be video. As a consequence, and for first time in the evolution

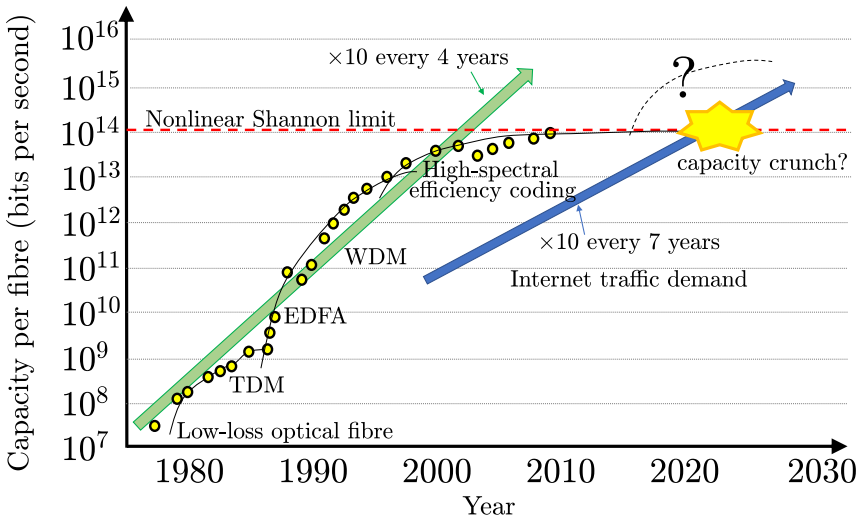


Figure 1.1 – Historical trend of the achieved capacity per fibre in bits per second versus the global Internet traffic demand, as well as the predicted capacity crunch within the next decade.

of data communications, it seems that we are facing a big issue: the so-called "capacity crunch" [4]. Predictions are not very optimistic and it is estimated that the bandwidth demand will overcome the expected available capacity in the next decade [5], which would surely result in the access to the Internet becoming a more expensive resource for everyone. To avoid such an undesirable scenario, a yet to arrive ground-breaking technology is yearned for and, in this direction, joint efforts are being made worldwide by the scientific community.

The global communications network is categorized in three levels according to their reach: **core networks** interconnect countries across long distances such as transoceanic links, reaching from hundreds to several thousands of kilometres, **metropolitan (or regional) networks** generally connect different cities within a country, ranging up to few hundreds of kilometres, and **access networks** (AN), also known as "last mile" networks, that connect operators and end users.

Although core and metropolitan networks benefit from great achievements such as recent transmission records from hundreds of Terabits per second (Tb/s) [6, 7] and up to several Petabits per second (Pb/s) [8, 9] through a single, multiple core, multi-mode fibre (practically disregarding how expensive or power hungry systems are), special attention is set upon the optical AN (OAN) because they are identified

as one of the bottlenecks, provided that cost and simplicity are key factors that operators require in order to adopt new solutions, posing an additional challenge: not only do they have to provide the demanded higher bandwidths but they also have to keep costs at its lowest in order to be commercially attractive and suitable for mass production.

As the part of the network responsible for the interconnection between the operator (also carrier or service provider) and its customers (also subscribers or users), which can be businesses or private households, the OAN consists of all the necessary equipment (including cables, optical fibres, and antennae) in order to provide connectivity between a central office (CO) equipped with an optical line terminal (OLT) and the end-user premises where the optical network unit (ONU) is located. Variants of OAN are referred to as "fibre to the X" (FTTx), where X can be office, cabinet, building, home, etc, depending on where the optical fibre terminates.

1.2 Current trends and enabling technologies

In order to keep up with the ever-increasing bandwidth demand, several technologies are envisaged as potential candidates to be incorporated in future communication networks. As stated above, the most sophisticated techniques usually find a place within the range of core and metropolitan networks despite their cost, which is shared amongst all the users of the network. In this regard, space division multiplexing (SDM) is gaining a lot of attention in recent literature [10–12], exploring the capabilities of multiple core fibres, or the independent use of different modes of propagation in few-mode, or multimode fibres (FMF or MMF, respectively).

On the other hand, the preferred approach for the deployment of OANs is that of *passive optical networks* (PONs) due to its low cost and simplicity [13]. PONs are the most deployed type of OAN. A basic PON consists of an OLT providing service to a number of ONUs via a passive network equipped with different types of optical fibre, optical passive splitters, and optical passive filters, along with all the interface connectors. Its main advantage is that none of these elements require a power supply. In order to benefit from these attractive properties of PONs while trying to increase the data throughtput per user, many research projects have seen

the light in recent years, aiming to identify the enabling technologies that could satisfy nowadays operator's requirements, and defining what was called the next generation optical access (NGOA) networks [14–21].

Coherent detection and digital signal processing

Coherent detection is probably one of the most consolidated and well-known techniques that readily bring a substantial improvement over its counterpart of direct detection [22–24]. The main benefit offered by coherent detection is the increased number of signalling dimensions, since the receiver not only has access to the signal amplitude but also to the phase. Other benefits of coherent detection are the improved sensitivity due to the amplifying effect produced by the beating with a local oscillator (LO) and the channel selectivity that can be achieved by accordingly tuning the LO centre wavelength.

Practically any optical coherent system is accompanied by further stages of digital signal processing (DSP) at the receiver that exploit the extra phase information obtained with the signal to compensate for transmission impairments [25–29]. The theoretical aspects of coherent detection and DSP are reviewed later in Chapter 2.

Coherent detection is ubiquitous in long-haul links and metropolitan networks, but its use has been historically absent in OAN due to the higher cost of integration of the required optical components, even though the idea of subscriber lines using this technology was proposed more than 30 years ago [30–33]. However, the concept of coherent OAN has a renewed interest, fostered by the maturity of semiconductor technologies with higher degree of integration at a lower cost. As a result, the last decade has witnessed a fair amount of research in this direction, hoping to consolidate the technology as the way to go in NGOA [34–39]. Additionally, a more recent standard known as 400ZR has been developed by the Optical Internet Forum (OIF) during the past few years, aimed to fulfil the need for increased bandwidth in data centre interconnects (DCI) within metropolitan networks. As such, it is defined for Ethernet-based single-fibre links transporting 400 Gbits/s, in a single-channel amplified DWDM architecture, and covering distances of up to 120 km [40, 41].

WDM and its flavours

Another technology that represented a huge breakthrough in optical communications is WDM. It is considered a media access (MA) technique, since it describes how different users get their data from a shared channel (the fibre). Its older counterpart is *time-division multiplexing* (TDM), where each user "sees" all the traffic at reception but it is only entitled to capture the content of one predefined time slot. In WDM, multiple wavelengths —coming from different laser sources or a single frequency comb— are independently modulated and multiplexed together within the same optical fibre. Therefore, each wavelength is associated to a transmission channel, usually targeted to a single receiver which filters out all but the selected wavelength, and hence does not have access to the rest of the channels. Since its early study and experimental verification applied to high-speed data transmission in the late 70s [42–44], it has become a fundamental part of global communications, specially in the long- and mid-range regimes, to the point that standards for coarse (CWDM) and dense (DWDM) grids have been proposed so far by the International Telecommunication Union (ITU), defining parameters such as frequency range and inter-channel spacing [45, 46]. More recently, other flavours of WDM using denser frequency grids purposely tailored to PONs, also called ultra-dense WDM (uD-WDM), have been investigated as a potential solution to provide multi-gigabit connections to a higher number of users [15, 47–51].

Directly modulated lasers

The use of directly modulated lasers (DML) is predominant in last-mile and short-range links such as data centres, typically under an intensity-modulation direct detection approach (IMDD), thanks to their low-cost and ease of manufacture. A comprehensive review of the historical evolution, types of lasers, and their applications in modern communication systems can be found in [52]. The inherent frequency chirp of semiconductor lasers have historically prevented them from being used in longer links (>20 km) due to the distortions produced when interacting with the fibre chromatic dispersion. Although early works proposed techniques to produce non-chirped pulses with a DML [53, 54], they do not seem to have gained the acceptance of the research community, maybe because the need of external op-

tical interferometers (implying insertion losses) or because the transmission of more than 1 Gb/s was not a priority at the time of publication (1988-1989). It has been in the last decade that performing some kind of phase modulation in DML has recovered more interest [55, 56], since it can provide a suitable solution for current and future cost and capacity demands in OANs.

Consequently, it must not be a surprise to see how in the last 5 years big efforts are invested towards the combination of these enabling technologies [57–63], promising multi-gigabit services to the increasing number of users while trying to continue being cost-wise competitive.

1.3 A need for improved channel approximations

According to *information theory* (IT), what prevents us from sending an infinite amount of information through a given channel is the presence of noise [64], which is random in nature. In an optical channel, noise comes from many and diversified sources. Having an exact model of the channel is what enables the optimum design of the different subsystems involved in order to maximise the capacity, but in high-speed communications the computation time is strongly restricted by the frequency of real-time operation of DSP circuits (on the order of several hundreds of MHz), thence, relying on approximations that alleviate the hardware complexity is a must. The most popular approximation is that of the *additive white Gaussian noise* (AWGN) channel, which considers all the noise sources in a lumped term of normally-distributed noise.

With the view set on the increasing number of proposed solutions that combine the technologies listed in Section 1.2, it appears necessary the implementation of alternative models that fix the weaknesses of the AWGN channel model, particularly the lack of sensitivity to angular or phase noise. This is so because, firstly, coherent detection relies on the beating of two oscillatory sources, rendering a signal that is corrupted by, not only AWGN, but also the phase noise that is intrinsic to all lasers [65]. Additionally, WDM introduces another undesired effect produced by the interaction of neighbouring channels, known as *cross-phase modulation* (XPM), that manifests itself as another source of phase noise [66–68], and that is expected to be

larger as the channels are closer together in wavelength. Lastly, the chirp produced by DMLs has also an angular nature, therefore any AWGN noise introduced while driving the laser will result in a chirp-induced phase noise contribution.

Moreover, the need for new models might be accentuated by recent trends suggesting new paradigms such as long-reach OANs [69–73] (such as the Super-PON concept, firstly introduced in 1996 [74], conceived to support split ratios of 1:1024 over 100-km links, and recently fostered by Google Fiber to economise the deployment of FTTH [75]) and elastic optical networks (EON) [76]. In the first case, the aim is to consolidate COs at the expense of longer links reaching up to 200 km or more, thus yielding a larger amount of accumulated chromatic dispersion and attenuation. The second case considers a network in which aspects like modulation format, bandwidth allocation, baud rate, or traffic routes are decided *on-the-fly* according to a set of monitored network parameters.

1.4 Thesis outline

After the brief introduction to the context of modern optical communications and the challenges faced by the research community that feed the motivation for endeavouring this work, the rest of the content of this document is structured as follows.

In Chapter 2, a basic theoretical background of the main components found in modern optical communication systems is provided. The chapter covers elements from the transmitter side, such as semiconductor lasers, modulator types and techniques as well as the most commonly used modulation formats. The optical fibre channel is briefly reviewed, highlighting its main transmission impairments. Lastly, receiver architectures and components are reviewed, paying special attention to coherent receivers and DSP compensation algorithms.

Chapter 3 is eminently dedicated to the modelling of the PCAWGN channel. Here, the exact solution for this type of channel is recovered from literature, and means to solve it in a computer avoiding numerical exceptions and overflows are provided. Beyond the exact solution, an alternative approximated model in polar coordinates is presented, yielding a lower-complexity expression that is shown to

approximate the channel with excellent accuracy. To complete the comparison, another previously published approximated model is also implemented and evaluated.

A collection of several applications that have been identified to be suited to use with the proposed model is given in Chapter 4. Some remarkable results are obtained in carrier phase estimation (CPE), where in combination with a decision-directed algorithm the model provides an increased tolerance to laser phase noise. Other explored applications are the definition of a slicing metric, symbol error rate estimation and blind estimation of the noise sources present in the channel.

A whole new topic (although still closely related to the PCAWGN channel), the complex modulation of directly modulated lasers (CM-DML), is investigated in Chapter 5. In this part of the work, a new way to approach the design of laser driving signals tailored to account for the chirp effect is proposed. Several modifications to previously published decoding algorithms are suggested and tested.

Lastly, a summary of the work presented in this thesis is found in Chapter 6, where further ideas are also suggested for a future research along the same conceptual lines given in this thesis.

1.5 Original contributions of this work

The main original contributions presented in this thesis are summarized in the following list:

- The most important contribution is the description of a reduced-complexity approximated model of the PCAWGN channel in polar coordinates, given in a compact expression that has been extensively evaluated and that is suited for multiple applications.
- A phase noise-sensitive metric is derived from the proposed approximation and is shown to be a suitable slicing function in the presence of phase noise.
- A method to modify the blind phase search algorithm (BPS), replacing the Euclidean distance by the proposed slicing metric, is shown to improve the linewidth tolerance of the carrier phase estimation (CPE) stage of a digital coherent receiver.

- A method to use the proposed model to perform hard-decision symbol error rate (HD-SER) estimation is also described, tested, and shown to exhibit an excellent agreement with Monte Carlo simulations.
- A method to perform blind joint estimation of SNR and phase noise variances from blocks of received, undecoded data, is presented and evaluated through simulations of multiple scenarios.
- Based on previous works on the complex modulation of directly modulated lasers (CM-DML), a new paradigm of constellation design, called the *spiral constellation* is described.
- Following the guidelines of this new paradigm, two different methods for the design of spiral constellations with DML are proposed, showing a big potential to improve further the prior SNR gain obtained in links provided with a coherent receiver.
- A modified version of the previously published Viterbi algorithm (VA) in CM-DML is shown to be further improved in the presence of non-negligible phase noise if the proposed polar metric is used as the transition distance.

Additionally, in the last sections of this thesis, two important contributions are introduced and, although they lack a substantial amount of work yet to be done, the preliminary results indicate that they are promising lines of future research:

- The proposed approximated model of the PCAWGN channel is shown to be readily upgradable to incorporate a nonlinear term, accounting for intensity-dependent phase contributions.
- An alternative version of the proposed model specifically modified to reflect a differential detection scheme is introduced. Early simulations show potential improvements in SNR gain of up to 2.2 dB when used with the VA in CM-DML.

1.5.1 Publications

Fruit of this work, two articles were submitted and accepted for publication:

- **Sales-Llopis, Martí**, Md Saifuddin Faruk, and Seb J. Savory. "Improved Linewidth Tolerant Carrier Phase Recovery Based on Polar MAP Metric Estimate". *In Optical Fiber Communication Conference*, paper Th4C.2. Optical Society of America, 2017.

- **Sales-Llopis, Martí**, and Seb J. Savory. "Approximating the Partially Coherent Additive White Gaussian Noise Channel in Polar Coordinates". *IEEE Photonics Technology Letters* 31, no. 11 (June 2019): 833–36.

- **Sales-Llopis, Martí**, and Seb J. Savory. "A Transition Metric in Polar Co-ordinates for MLSE of a Complex Modulated DML". *In Optical Fiber Communication Conference*, paper M3J.6. Optical Society of America, 2020.

Chapter 2

Theoretical background

This chapter is intended to provide the background on the physical and functional aspects of an optical communication system. The goal is to familiarise the reader with the underlying concepts that constitute the basis for the research presented throughout this thesis. Virtually any communication system can be represented by the block diagram shown in Figure 2.1 [64]. In principle, the topic of source coding and decoding —corresponding to the first and last blocks— falls beyond the focus of this work and, unless stated otherwise, in the study cases considered herein it will be assumed that the data to be transmitted has been already encoded. That said, and following the flow shown in Figure 2.1, this chapter is structured as follows: Section 2.1 covers the transmission side of the optical link, describing the most commonly used optical sources available and some of the different modulation devices and techniques. Section 2.2 describes the channel medium: the optical fibre, briefly exposing its main physical characteristics, then introducing the two main sources of linear impairments, namely attenuation and chromatic dispersion. Section 2.3 will expose the two main receiver architectures, direct and coherent detection.

For the purposes of this work, special attention has to be paid to the latter case, in which the system has access to both amplitude and phase of the received signal. This leads to the fact that the receiver not only has to deal with noise contributions to the signal amplitude but also to its phase, which usually introduces memory in the channel. Hence, the effects of phase noise and the problem of carrier phase

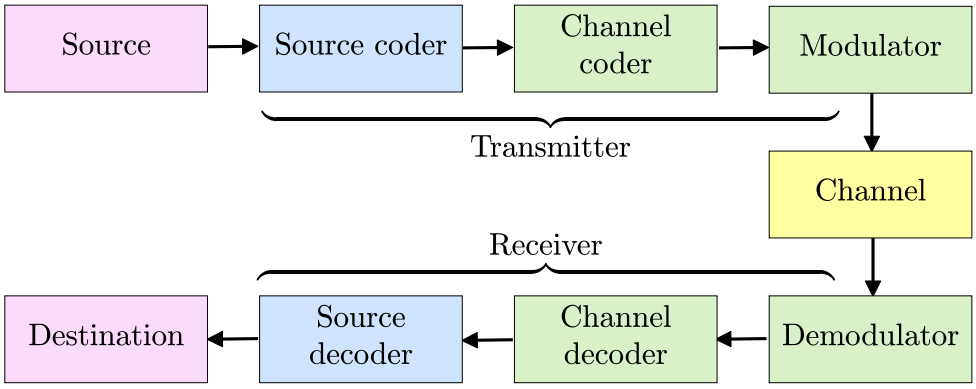


Figure 2.1 – Block diagram with the typical elements in a communication channel. This thesis is mainly related to the boxes on the right half, in green and yellow colour: channel de/coding, de/modulation, and optical fibre channel.

estimation will be studied in Section 2.4, a technique from which the model of the so-called *partially coherent* channel is derived and that will be the main subject of research in the next chapters.

2.1 Optical transmitters

The purpose of the transmitter in an optical communication system is to generate an optical signal by means of a light source and a modulating electrical signal that conveys the encoded user data, and send it through an optical fibre cable that will reach the receiver at the other end. In this section, the main types of light sources based on semiconductor devices are reviewed, along with the two basic ways of optical modulation, namely, internal and external. Also a survey of simple and advanced modulation formats is presented, from the basic on-off keying (OOK) to high order quadrature amplitude modulation (QAM) constellations.

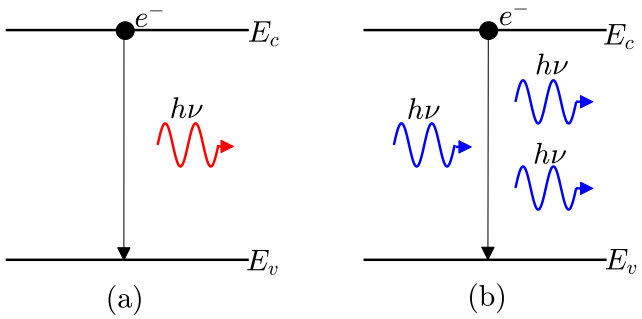


Figure 2.2 – Photon radiation mechanisms: spontaneous (a) and stimulated (b) emission. Both require an electron falling from the conduction band, with energy E_c, to the valence band, with energy E_v, releasing a photon with energy $E_{ph} = h\nu$.

2.1.1 Semiconductor lasers

Principle of operation

Semiconductor devices such as light-emitting diodes (LED) or *lasers* are by far the most employed light-generation devices that rely on electrical driving signals. Both of them are, in its most basic form, PN-junction diodes that operate thanks to the electroluminescence effect, by which excited electrons in the conduction band recombine, decaying to the valence band, and releasing energy in the form of photons. The wavelength of these photons is related to their energy, E_{ph}, by means of Planck's constant h, such that

$$E_{ph} = h\nu = h\frac{c}{\lambda_{ph}} \tag{2.1}$$

where ν is the optical frequency in Hertz, c is the speed of light in vacuum and λ_{ph} is the photon wavelength.

The main difference between LED and lasers is the dominant radiation mechanism. In the former case, *spontaneous* radiation is the dominant mechanism and, as such, the photons generated are more diverse in terms of emitted wavelength and phase, thus emitting light that presents lower coherence or, in other words, a higher linewidth. Therefore, while a LED modulation bandwidth can be between a few hundred of MHz and up to 1 GHz (since it is governed by the carrier lifetime in the order of 1-100 ns) its large emitted-light spectral bandwidth (in the order of tens of

terahertz) deprives them from being practical in fibre optics communications due to dispersive nature of the fibre. They might be, however, suitable for small in-house or in-car networks that operate at low data rates and with reaches spanning a few metres. The concept of spontaneous emission is illustrated in Figure 2.2 (a).

On the other hand, lasers –which name actually stands for *light amplification by stimulated emission radiation*— rely on the generation of photons that have been stimulated by another passing-by photon, process illustrated in Figure 2.2 (b).

The most remarkable characteristic of the stimulated emission process is that a stimulated photon is identical to the originating one in frequency, phase, and direction. This, in turn, leads to a closer approach to the generation of monochromatic light, with linewidth, power and stability characteristics depending on the laser structure.

The first and simplest laser structure is known as the Fabry-Perot laser, consisting of an optical cavity terminated by cleaved facets that act as reflectors. Its purpose is to create resonance conditions inside the cavity so photons are reflected back and forth by the mirrors, and hence stimulating the generation of new photons.

Other more sophisticated laser types are the distributed feedback (DFB) and distributed Bragg reflector (DBR) lasers, in which the optical feedback is provided by means of a corrugated structure (inside or outside the gain medium, respectively) that presents longitudinal periodicity in its refractive index, helping towards a more stable single-mode operation and reduced linewidth [77].

Rate equations

The single-mode time evolution of photon and carrier densities ($S(t)$ and $n(t)$, respectively) in the laser cavity is commonly described by the pair of coupled differential equations [78]:

$$\frac{dn}{dt} = \frac{I_A}{qV_{act}} - \frac{n}{\tau_n} - g_0 (n - n_0)(1 - \epsilon S) S \tag{2.2}$$

$$\frac{dS}{dt} = \left\{ \Gamma g_0 (n - n_0)(1 - \epsilon S) - \frac{1}{\tau_p} \right\} S + \Gamma \beta \frac{n}{\tau_n} \tag{2.3}$$

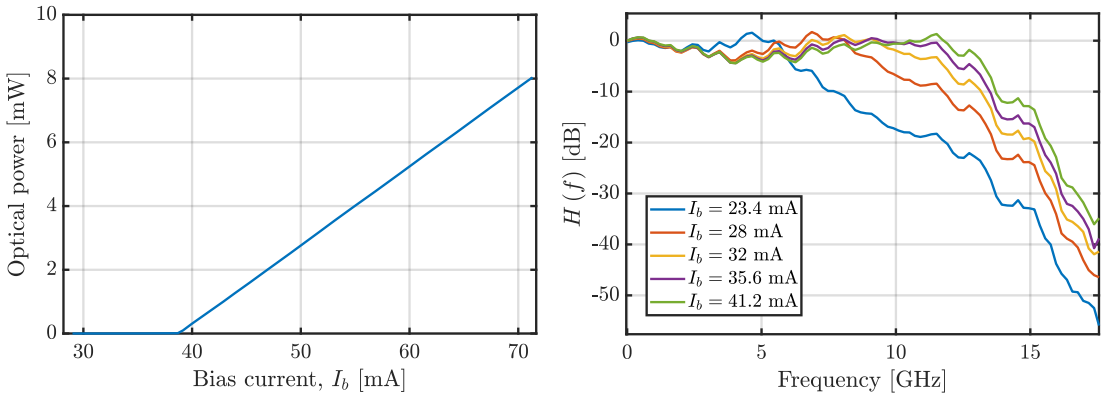


Figure 2.3 – Experimentally obtained light-current (left) and frequency response for different bias (right) curves for a DFB laser with 10-GHz nominal analogue bandwidth.

where Γ is the optical confinement factor, g_0 is the gain slope constant, n_0 is the electron density at which the net gain is zero, τ_p is the photon lifetime, τ_n is the electron lifetime, β is the fraction of spontaneous emission coupled into the lasing mode, V_{act} is the volume of the active layer, q is the electronic charge, I_A is the current injected into the active layer, ϵ is the gain compression parameter. In addition, the output optical power is related to the photon number S as

$$P = \eta \frac{Shc}{\lambda} \tag{2.4}$$

where η is the *quantum efficiency* of the laser. Two important results that can be obtained from the analysis of the rate equations are the light-current curve and the frequency response of the laser. However, some of the parameters are not often known beforehand and need to be extracted from experimentally measured data [79]. As an example, Figure 2.3 shows the measured light-current of a DFB laser (left) and its frequency response for five different values of bias current (right).

The equation to obtain the frequency chirp can be obtained directly from the output optical power:

$$\Delta f = \frac{\alpha}{4\pi} \left(\frac{d}{dt} \ln P(t) + \kappa \cdot P(t) \right) \tag{2.5}$$

in Hz, where $P(t)$ is the output optical power, α is the linewidth enhancement

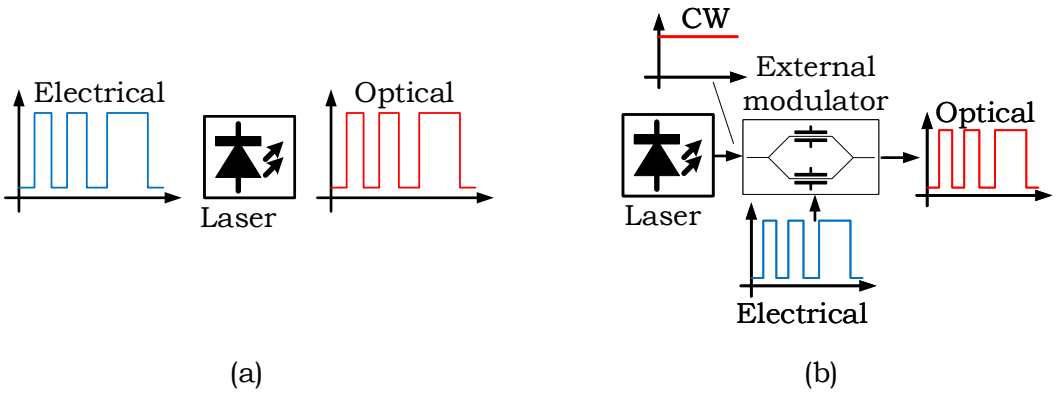


Figure 2.4 – Direct modulation vs External Modulation

factor, and κ is the adiabatic chirp coefficient.

2.1.2 Optical modulation techniques

Optical modulation is the process by which one or more electrical signals will shape an optical signal to convey the transmitted data. The parts of the optical signal that can be modulated are its amplitude (or power), its frequency, phase, polarisation, or a combination of these. Figure 2.4 shows two different paradigms of optical modulation: direct and external, both briefly described within this section.

Direct modulation

Direct modulation is referred to the technique by which the optical signal amplitude is modified by means of the laser's driving current. Under this mode of operation, a laser is commonly referred to as a directly modulated laser (DML). The simplest and most widely deployed technique of this kind is the *intensity modulation* (IM), from which derives the popular OOK modulation format, by which digital data is imprinted on the optical signal by injecting low and high values of driving current, which produce two levels of output optical power, i.e. low ("0") and high ("1"), respectively. When using this approach it is common to bias the laser so it is centred within its linear range (see Figure 2.3 left), with the low value

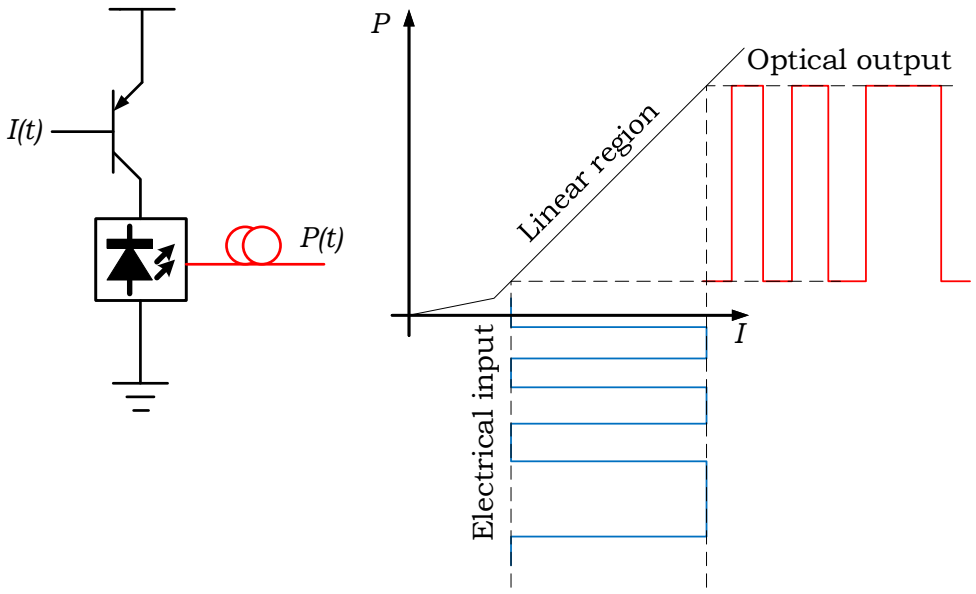


Figure 2.5 – Simple schematic of a OOK generating circuit by intensity modulation (IM)

for current being slightly above threshold. A simple representation of an IM circuit and the input-output curves is shown in Figure 2.5.

A variation of direct modulation that has received great interest in the recent years is the so-called *complex modulation* (CM) [80]. This technique exploits the property of the chirp effect in a directly modulated laser (given by Equation (2.5)) by which the optical signal will suffer a phase shift proportional to the change in its intensity. Using a coherent receiver, and although there is only one degree of freedom at the transmitter side (i.e. the eletrical drive amplitude), the receiver can obtain a significant amount of information from the observed phase, being able to substantially reduce the signal-to-noise ratio (SNR) requirements for a given BER target. This topic will be covered in more detail in Chapter 5.

External modulation

External modulation refers to the technique by which the optical source produces a continuous wave (CW) generated by a constant injection current (only adjusted in a slow manner to compensate for temperature-induced drifts) while another device

will operate upon the incoming CW signal. In modern systems we find basically two types: electro-absorption modulators and Mach-Zehnder modulators.

Electro-absorption modulator: an EAM consists of a semiconductor device that uses an electric field generated by an externally applied voltage to change the absorption spectrum of light of a waveguide subject to the field. This principle, known ad the Franz-Keldysh effect [81], is used to modulate the intensity of an optical signal. The main advantages of EAM are its small size, low driving voltage and high bandwidth. The smaller size property makes it a suitable candidate for high-density integration, e.g. sharing a package with DFB lasers, which reduces insertion losses. The lower voltage requirements make it suitable for low-power devices and relaxes the driving electronics requirements. Finally, the higher bandwidth of up to 100 GHz [82], yields EAM an attractive solution for high-speed telecommunications. The main disadvantage is the higher chirp effect compared to MZM, which introduces severe inter-symbol interference (ISI) when significant amounts of dispersion are introduced in the channel. However, it has been investigated and argued the feasibility of exploiting the chirp in frequency shift keyed (FSK) modulation [83].

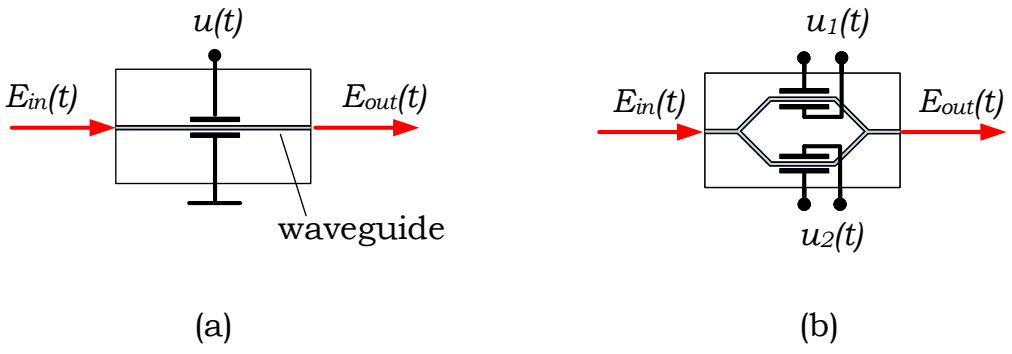


Figure 2.6 – Strucure of the (a) phase and (b) the Mach-Zehnder modulators

The Mach-Zehnder modulator To achieve higher-order modulation formats (as those reviewed in the forthcoming Section 2.1.4) it is necessary to independently modulate the amplitude and the phase of the signal. A basic component for this

purpose is the phase modulator (PM). A PM is similar to an EAM in the sense that it relies on an induced electric field to change a property of the waveguide material, but in this case the changing property is the refractive index (which is known as the *Pockels effect*, since it was first described by the German physicist Friedrich Pockels in 1906). Thus, by means of an applied voltage $u(t)$ it is possible to change the phase of the incoming optical signal such that

$$E_{out}(t) = E_{in}(t) \exp\left(j \frac{u(t)}{V_\pi}\right) \tag{2.6}$$

where $E_{in}(t)$ and $E_{out}(t)$ are the input and output electric fields of the optical signal, $u(t)$ is the applied voltage, and V_π is a design parameter representing the necessary voltage to induce a phase shift of π radians (180°). Equation (2.6) neglects any insertion loss. Two PMs can be coherently combined to form what is known as the Mach-Zehnder modulator (MZM), which is a device that equally splits the input signal into two paths, each of which goes through a different PM applying relative phase shifts φ_1 and φ_2, and then combines these signals again. This way it is possible to also modulate the optical intensity of the resulting signal. In a push-push configuration (i.e. $\frac{u_1(t)}{V_{\pi,1}} = \frac{u_2(t)}{V_{\pi,2}}$), and the modulation is entirely produced in phase, whereas in a push-pull configuration $\varphi_1 = -\varphi_2$ (i.e. $\frac{u_1(t)}{V_{\pi,1}} = -\frac{u_2(t)}{V_{\pi,2}}$) and the modulation is produced entirely in amplitude. The electric field at the output of a MZM can be expressed as [84, Ch. 2.1.2]

$$E_{out}(t) = \frac{E_{in}(t)}{\sqrt{2}} \left[Y_1 e^{j\left(\frac{u(t)}{V_\pi}\right)\pi} + Y_2 e^{js\left(\frac{u(t)}{V_\pi}\right)\pi} \right] \tag{2.7}$$

where we distinguish two important modes of operation (ideally assuming $Y_1 = Y_2 = \sqrt{0.5}$): the case $s = 1$ corresponds to the "push-push" configuration where $\varphi_1 = \varphi_2$, and the modulation is produced entirely in the phase. Thus, the output electric field is given by

$$E_{out} = E_{in} \exp\left(j \frac{u(t)}{V_\pi} \pi\right) \tag{2.8}$$

while the case $s = -1$ corresponds to the "push-pull" configuration where $\varphi_1 = -\varphi_2$, and the modulation is produced entirely in the amplitude with the additional benefit

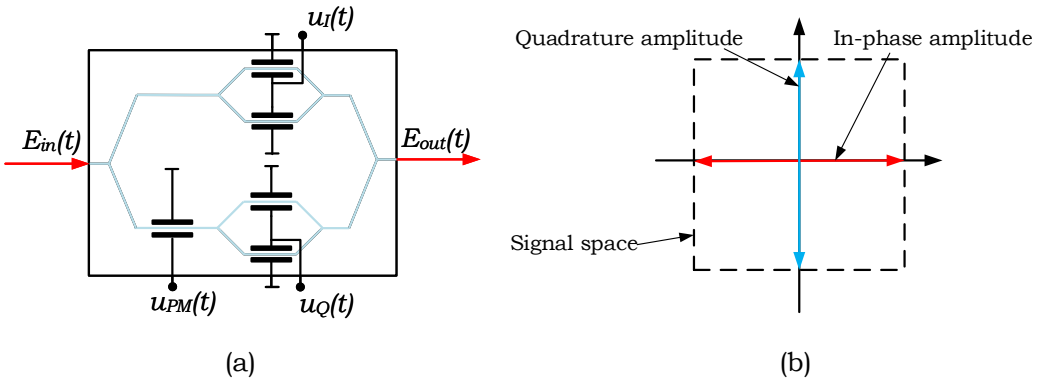


Figure 2.7 – (a) Structure of the in-phase and quadrature modulator (IQM), with two Mach-Zehnder modulators (MZM) and one phase modulator (PM); (b) Signal space at the output of the IQM.

of being free of frequency chirp. Here the output electric field is obtained as

$$E_{out} = E_{in} \cos\left(\frac{u\,(t)}{V_\pi}\pi\right) \tag{2.9}$$

More generally, $Y_1 = \sqrt{0.5+\epsilon}$ and $Y_2 = \sqrt{0.5-\epsilon}$, where ϵ represents the deviation from perfect splitting (i.e. when $\epsilon = 0$). The ratio of the maximum to minimum output power is known as extinction ratio (ER) and is calculated as

$$ER = \frac{(Y_1 + Y_2)^2}{(Y_1 - Y_2)^2} \tag{2.10}$$

where insertion losses have been neglected. Additionally, one way of achieving simultaneous and independent modulation of amplitude and phase is by employing the structure given in Figure 2.7 (a) known as in-phase and quadrature modulator (IQM). In a IQM, the input signal is split into two halves and each is directed to an independent MZM, although one of them has been first phase-shifted by $\frac{\pi}{2}$ radians, i.e. setting $u_{PM}\,(t) = -V_\pi/2$. Finally, the two modulated halves are summed to form the output signal, whose signal space is represented in Figure 2.7 (b).

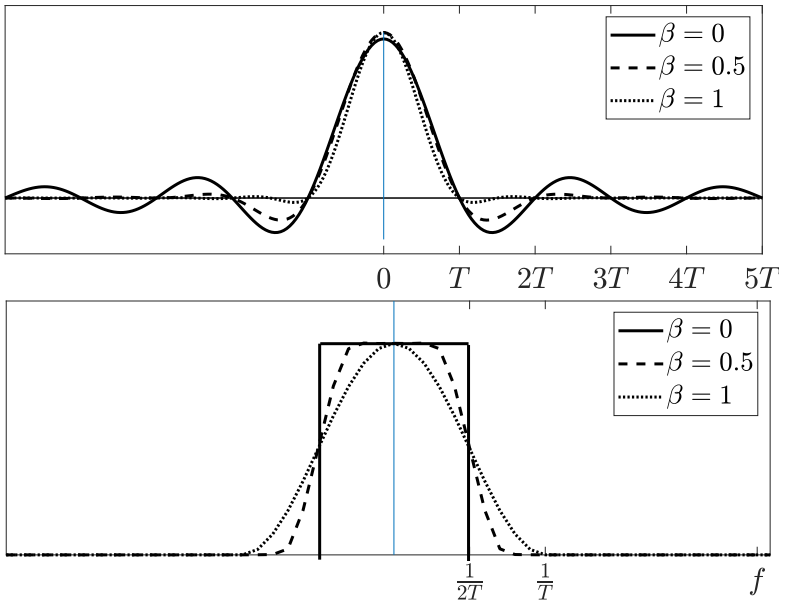


Figure 2.8 – Raised cosine filter: time evolution (top) and frequency response (bottom) for three values of β, equal to 0 (solid), 0.5 (dashed), and 1 (dotted).

2.1.3 Pulse shaping

Transmission channels have a finite bandwidth and their spectrum may also be spliced into a number of sub-channels of limited frequency range in which the user is allowed to transmit. If transmitted pulses were to be square wave signals representing data, this would violate the bandwidth restrictions because such signal contains an infinite amount of odd harmonics beyond the main frequency.

Therefore, a technique called pulse shaping (PS) is put into practice, consisting of shaping the waveform of the transmitted pulses to limit their bandwidth and hence to better comply with the channel specifications and its spectrum allowance. In optical systems, the PS process is generally applied to the driving electric signal, either in direct or external modulation of lasers. Digital filters (such as *finite impulse response*, FIR) followed by digital-to-analogue (DAC) converters are the most common way to implement PS filters.

The ideal PS filter is known as the *boxcar* filter, since it presents a rectangular shape in the frequency domain. In the time domain, the boxcar filter impulse

response is obtained as

$$h(t) = \text{sinc}(t) = \frac{\sin(t)}{t} \tag{2.11}$$

However, the boxcar filter is impractical in real systems because of causality issues and its infinite length in time. A good approximation and popular choice in practical implementations is the *raised cosine* filter (RC). As its name indicates, the spectrum of the RC filter has a cosine shape that is raised to sit on the zero line. What makes the RC filter attractive is not only that it is realizable with digital filters but it complies with the Nyquist ISI criterion (the contribution of its response to integer multiples of T is zero except for $t = 0$) while offering a trade off between hardware complexity and performance. The frequency spectrum of the RC filter is calculated as:

$$H_{RC}(f) = \begin{cases} T & \left(0 \leq |f| \leq \frac{1-\beta}{2T}\right) \\ \frac{T}{2}\left\{1 + \cos\left[\frac{\pi T}{\beta}\left(|f| - \frac{1-\beta}{2T}\right)\right]\right\} & \left(\frac{1-\beta}{2T} \leq |f| \leq \frac{1+\beta}{2T}\right) \\ 0 & \left(|f| > \frac{1+\beta}{2T}\right) \end{cases} \tag{2.12}$$

where β is the *roll-off* factor and indicates the *excess bandwidth* with respect to the ideal $\frac{1}{T}$. Figure 2.8 shows the time (top) and the frequency (bottom) responses of a RC filter for the β values 0, 0.5, and 1.

In the presence of noise, it is common to exploit the properties of the autocorrelation function in order to reduce out-of-band noise power. This technique is referred to as *matched filtering* and consists of designing the transmitter and receiver filter such that the convolution between both yields the desired channel response. In this regard, the *root raised cosine* (RRC) filter is designed as the square root of the RC filter, that is $|H_{rrc}(f)| = \sqrt{H_{rc}(f)}$, and is implemented in both sides of the channel so we can say that transmitter and receiver are *matched*.

2.1.4 High order modulation formats

As customer needs evolved alongside increased bandwidth demand, digital communications systems needed to improve the existing solutions offering higher spectral efficiency, that is, transmitting more data bits per second per Hertz. The best way to achieve this goal is using multilevel signalling, which consists of increasing

the number of discrete levels that a receiver will be able to distinguish with a relatively low error rate. Thus, the N-dimensional signal space is distributed in a set of M symbols that form the modulation format constellation. We will call *alphabet* or set of symbols $\mathcal{X}$, and individual symbols within $\mathcal{X}$ will be noted with a lower-case x_i where i will be the symbol index, i.e. $i \in \{1, \ldots, M\}$. In general —as in non-coded modulation— each symbol x_i will represent a unique codeword of $m = \log_2(M)$ bits, so that we can use the bit vector representation $x_i := [b_0, b_1, \ldots, b_m]$.

Pulse amplitude modulation

In IM systems, such as a transmitter based on a DML, increasing the number of intensity levels leads directly to *pulse amplitude modulation* (PAM). When using PAM modulation, the receiver can be as simple as requiring only one photodiode, thus becoming an attractive solution for access networks and last-mile topologies where cost has to be minimized. To this date, a substantial amount of research has been done for PAM4 and PAM8 reaching transmission rates of up to 200 Gb/s [85, 86].

Figure 2.9 depicts an example for PAM4 modulation including the signal space in the IQ plane (left) and a modulated signal after being PS filtered (right). As it will be seen in Chapter 5, if a laser has been directly modulated with a PAM signal, it is possible to highly improve the link performance if a coherent receiver is used due to the chirp present in the signal and thanks to the retrieved phase information.

Quadrature amplitude modulation

The most popular high-order modulation format found in industry is the multilevel *quadrature amplitude modulation* M-QAM. These, can be implemented in dual polarisation (DP) or single polarisation (SP). For simplicity, we consider the latter for the general description of the concept. M-QAM is achieved by modulating two carriers of the same frequency but in quadrature ($\pi/2$ phase shifted) with $\sqrt{M}$ amplitude levels and then combining them in the same signal. For instance, a SP-16QAM optical signal can be achieved by inputting a CW into the IQM mod-

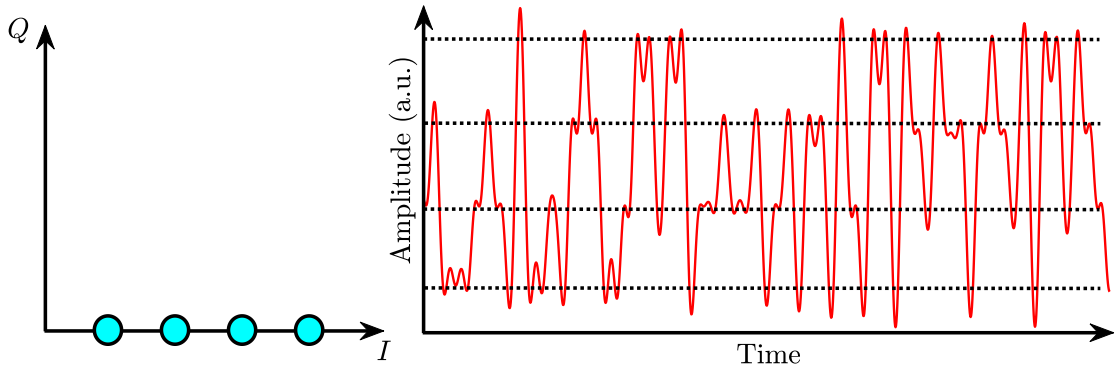


Figure 2.9 – PAM4 modulation: (left) signal space in the IQ plane, (right) modulated and pulse-shaped signal example.

ulator in Figure 2.7 and driving each MZM, i.e. the corresponding to I and Q components, with a PAM4 signal each.

The simplest QAM is the one with $M = 4$, equivalent to a *phase shift keyed* (PSK) modulation with 4 phase angles (4-PSK), most commonly known as *quadrature phase shift keying* or QPSK. It is certainly the most widely deployed modulation format in modern optical communications although denser formats of orders from 8 to 64 are also being intensely investigated and utilised in an increasing number of systems. Figure 2.10 shows some examples for different QAM modulation formats, where for the particular case of 8-QAM the circular configuration has been chosen against the rectangular one due to its higher robustness against phase noise).

QAM orders beyond 64 have been investigated with 4096-QAM format experimentally demonstrated over a 160-km link [87]. Due to its relevance, a great part of the work in this thesis is centred around this type of modulation format.

2.2 Optical fibre channel

The huge success of optical communication systems greatly rely on groundbreaking advances in optical fibre technology that started back in 1950 with the addition of a cladding layer but really triggering its industrial expansion thanks to the important contribution in 1979 with the reduction of the fibre attenuation down to 0.2 dB/km [88]. The aim of this section is to summarize the main characteristics

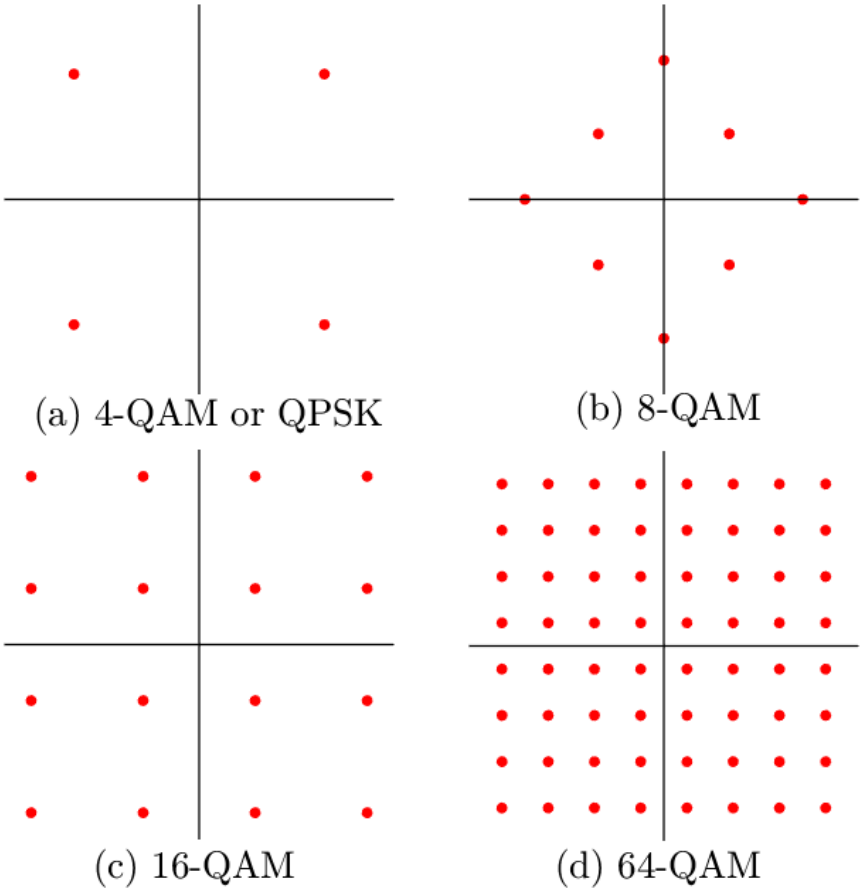


Figure 2.10 – Example of some QAM orders

of the fibre optic cable along with its main transmission linear impairments, with the view set on their implication in coherent systems.

2.2.1 Main characteristics

The optical fibre consists of a cylindrical strand of silica in its core surrounded by a cladding with lower refractive index, so that *total internal reflection* happens at the core-cladding interface and the light beam is guided within the fibre. Here we consider the type in which the change in refractive index between the core, n_1, and the cladding, n_2, is produced in a sudden step fashion (see Figure 2.11). This is known as the *step-index* fibre.

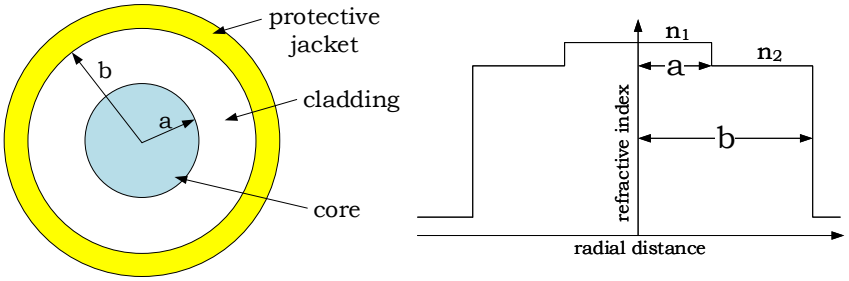


Figure 2.11 – Step-index fibre: cross-section view of the physical layers (left) and refractive index diagram showing the step between core and cladding (right).

Moreover, the core and cladding refractive indices and the radius of the core, a, will also determine the number of modes allowed in the fibre, that is, the number of propagation modes that are solutions to the Maxwell equations governing the wave propagation in the core-cladding medium. In this thesis, only single-mode fibre is considered, for which typical values are $\Delta = (n_1 - n_2)/n_1 \sim 3 \times 10^{-3}$ and radius in the range of $2 - 4\mu$m.

2.2.2 Linear impairments

Attenuation

Optical losses are one of the main impairments that limit the maximum distance a system can transmit in order to meet a given set of target specifications. For a fibre loss coefficient α, the received power P_{rx} at the other end of a fibre of length L (in metres) is given by

$$P_{rx} = P_{tx} \exp(-\alpha L) \tag{2.13}$$

where P_{tx} is the power launched into the fibre. α has units of *nepers*/m, however it is common to do the following conversion to units of dB/m.

$$\alpha_{dB} = -\frac{10}{L} \log_{10}\left(\frac{P_{rx}}{P_{tx}}\right) \tag{2.14}$$

The fibre loss is wavelength dependent, as it can be observed in Figure 2.12, presenting a minimum at the wavelength 1.55 μm (in the C-band), where most long-haul communication systems operate and which coincides with the window of effect of

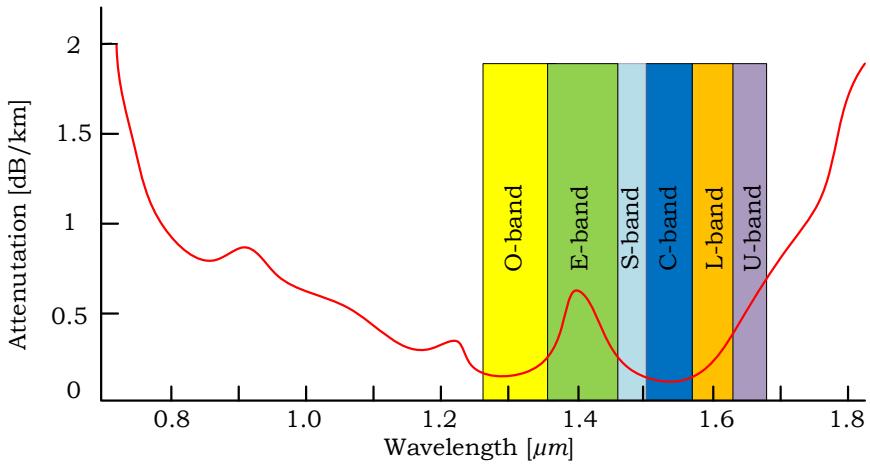


Figure 2.12 – Attenuation in dB/km versus the optical wavelength in μm, with the standard operation bands indicated.

erbium-doped fibre amplifiers (EDFA).

Dispersion

Dispersion refers to the phenomenon by which a pulse broadens as it travels along the optical fibre. Thus, a sent optical pulse of duration T will be received at the fibre output with a width $T + \Delta T$.

Depending on the size of the fibre core, there are two types of dispersion. Optical fibres allowing more than one mode (e.g. MMF) present *intermodal* dispersion, which is due to the different effective length that each mode has to travel inside the fibre. If the fibre core has a constant refractive index, then these mode components travel at the same speed, resulting in a difference in time of arrival at the output. Intermodal dispersion can be highly minimized by using *graded-index* fibres, a more sophisticated type of fibre in which the core refractive index decreases gradually and radially outwards from the centre, hence light rays travelling longer distances will do it faster, effectively compensating for the dispersion effect. Moreover, intermodal dispersion can be completely eliminated by using SMF.

Another two important types of dispersion present in optical fibres are *chromatic* dispersion (CD) and *polarisation mode dispersion* (PMD). On one hand, CD is the combination of two distinct effects: *material* and *waveguide* dispersion. The

former arises from the variable response at atomic scale of the interaction between dielectric and the electromagnetic field of the signal, translating into a wavelength dependence of the refractive index seen by the signal; the second also exhibits a wavelength dependence of the mode propagation velocity, but irrespective of any refractive index variations in the medium. On the other hand, PMD results in the two orthogonal states of polarisation (SOP) travelling at different speeds due to imperfections, asymmetry, or *birefringence* in the fibre core (that is, the difference of refractive index seen by these two SOP). For a more detailed description of these phenomena, the reader is addressed to [89]. Consequently, because optical pulses will contain a finite amount of frequency components, each one will travel at a slightly different group velocity, v_g, causing the so-called *group velocity dispersion* (GVD):

$$v_g = \frac{c}{n_g(\omega)} \tag{2.15}$$

where n_g is the group index [90]. For a pulse with frequency content defined by $\Delta\lambda$ after propagation of a distance L, the total dispersion is given by

$$\Delta T = DL\Delta\lambda \tag{2.16}$$

where D is known as the *dispersion parameter* and is commonly given by units of ps/(km·nm). It can be obtained as

$$D = -\frac{2\pi c}{\lambda^2}\beta_2 \tag{2.17}$$

where β_2 is the GVD coefficient and given by

$$\beta_2 = \frac{d^2\beta}{d\omega^2} = \frac{d^2(\overline{n}\omega/c)}{d\omega^2} \tag{2.18}$$

In Equation (2.17), higher-order effects have been omitted (such as β_3, also known as the dispersion slope). A typical value for standard SMF (SSMF) is $D \sim 17$ ps/(km·nm). As with attenuation, the amount of dispersion accumulated along the transmission link limits the reach and the data rate. It is one of the major impairments in mid-range long-haul transmission links but thanks to being a linear effect it is relatively easy to compensate for. Section 2.3.3 discusses how this can be done in the digital domain.

2.3 Optical receivers

The goal of an optical receiver is to collect the light that arrives from the fibre output, convert it to the electrical domain, and extract the information conveyed by the signal. There are two main schemes in which receivers can be categorized: *direct* detection and *coherent* detection. Whilst the former has been historically the dominant solution in industry, the latter is nowadays regarded by operators as a must-be solution in modern, high-bandwidth systems, not only in long-haul and metropolitan networks but also in short-range access networks. The two approaches are reviewed in Section 2.3.1 and Section 2.3.2, respectively. However, the aim of this thesis is particularly focused on coherent detection and, as such, more detail will be provided in this regard.

2.3.1 Direct detection

The simplest form of optical communication is given by an intensity modulated signal at the transmitter and a direct detection at the receiver. This scheme is abbreviated as IMDD. The receiver consists basically of a photodiode that retrieves the envelope of the incoming signal, where the information has been modulated. In such scenario, digital bits are encoded as low ('0') or high ('1') values of optical intensity at a modulation frequency many orders of magnitude lower than the optical carrier frequency (tens of GHz versus hundreds of THz). The photodiode is sometimes referred to as a *square-law* detector because the generated electrical current, I_{pd}, depends linearly on the incident electric field squared, that is

$$I_{pd} = R_{pd} \left| E_{in} \right|^2 = R_{pd} \left(\Re \left\{ E_{in} \right\}^2 + \Im \left\{ E_{in} \right\}^2 \right) \tag{2.19}$$

where R_{pd} is the photodiode's responsivity, obtained as

$$R = \eta \frac{q}{h f_{in}} \tag{2.20}$$

in units of ampere per Watt, where q is the charge of an electron, η is the quantum efficiency, h is Planck's constant, and f_{in} is the optical frequency of the incident light. An interpretation of Equation (2.20) is that it is telling us how many electrons

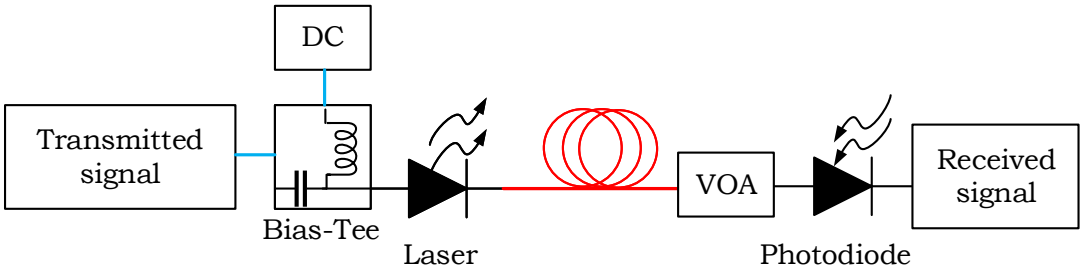


Figure 2.13 – A basic block diagram of the intensity-modulated direct detection approach (IMDD). VOA: variable optical attenuator.

(q) are generated per incident photon (each with energy hf_{in}), all being reduced by a factor that measures the "quality" of the device (normally $0 < \eta < 1$). Responsivity strongly depends on the incident wavelength, but typical values for devices operating around 1550 nm are around $0.5 - 0.6$ A/W. Figure 2.13 shows a basic block diagram of a IMDD link where, at the transmitter, the bias-tee allows for adding a DC bias current and the modulating RF signal whereas, at the receiver, the incident power can be adjusted with a variable optical attenuator (VOA). Alternatively, the intensity modulation can be performed externally with components such as MZM offering higher bandwidth. Using the latter approach, a record net data rate of 333 Gbps has been recently reported in [91].

2.3.2 Coherent detection

As with many other techniques used in optical systems, coherent detection [24, 92] is a technique borrowed from the field of electronic communications, and which main principle is the fact that combining (or *beating*) an incoming optical signal with a reference signal generated locally results in a *down-converted* version of the incoming signal. Unlike IMDD systems, coherent detection allows for recovering both phase and amplitude information from the modulated signal because the incoming signal is first split in two and then each part multiplied by the in-phase and quadrature components of the reference signal, respectively. An optical coherent receiver consists mainly of an optical front-end where the optical beating is produced, and the opto-electronic conversion block yielding the set of electrical RF

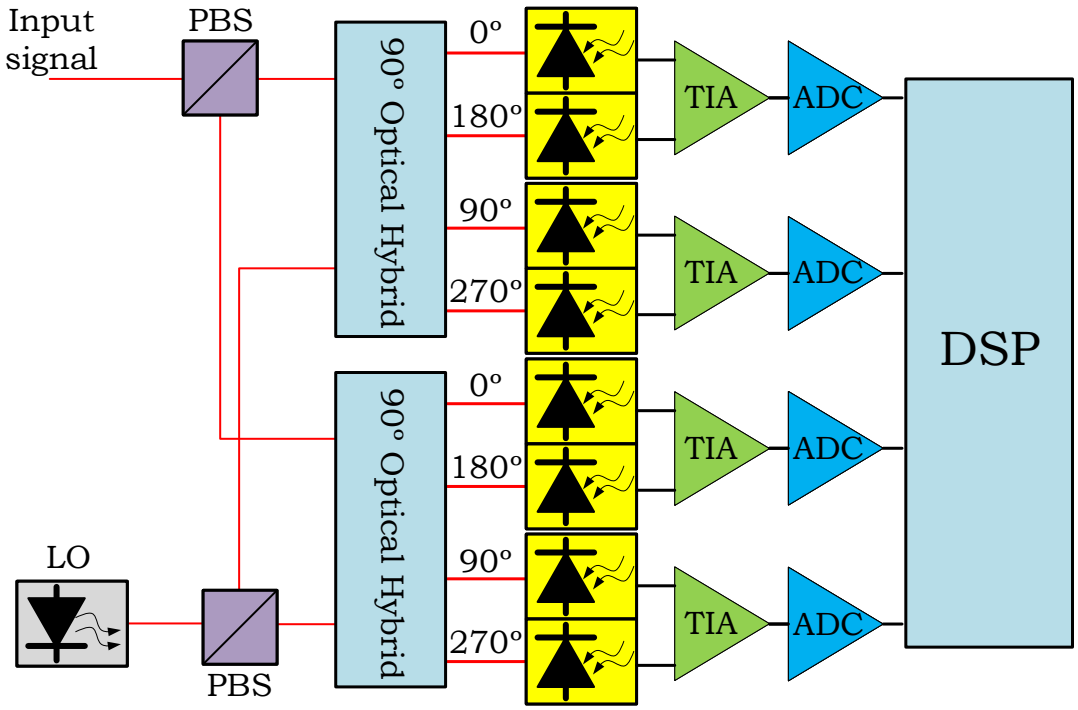


Figure 2.14 – Schematic of a phase- and polarisation-diverse optical coherent receiver. PBS: polarisation beam splitter, LO: local oscillator, TIA: trans-impedance amplifier, ADC: analogue-to-digital converter, DSP: digital signal processing.

signals. Moreover, since the mid 2000's until today, it has become extremely popular to exploit the power of *digital signal processing* (DSP) on the coherently detected signals, allowing for low-cost and reliable mitigation of transmission impairments such as CD, among others [25, 26]. In the remainder of this section the optical front-end architecture is briefly discussed along with the optoelectronic conversion stage. Some of the most important DSP compensation techniques are separately discussed in Section 2.3.3.

Optical front-end

This is the first block of an optical coherent receiver and it is composed by all-optical passive elements such as power splitters, couplers, etc. The number and configuration of these depends on the specific type of receiver, namely: whether it presents polarisation- and/or phase-diversity, single or balanced detection, etc.

For the sake of completeness, a phase- and polarisation-diversity coherent receiver will be herein described, although for most of this work a single polarisation is only considered.

Figure 2.14 shows the schematic of a phase- and polarisation-diverse coherent receiver, however, for the remainder of this thesis, only the part for one polarisation will be addressed, since the aspects that will be discussed are equally and independently applicable to both polarisations. There are two optical inputs for the signal and local oscillator (LO), whose electric fields can be expressed as

$$E_{in} = \sqrt{P_{in}(t)} \exp\left(2\pi f_{in} t + \theta_{in}(t)\right) \tag{2.21}$$

$$E_{LO} = \sqrt{P_{LO}(t)} \exp\left(2\pi f_{LO} t + \theta_{LO}(t)\right) \tag{2.22}$$

where $P_{in}(t)$ and $P_{LO}(t)$ are the optical power, f_{in} and f_{LO} the carrier centre frequencies, and $\theta_{in}(t)$ and $\theta_{LO}(t)$ the phase of the incoming and LO signals, respectively. These two signals are fed into an optical hybrid that contains two 2×2 3-dB couplers. These couplers receive half of the two signals at its outputs, with the difference that for one of them the LO signal has been previously shifted by 90°. Each coupler adds a 180° shift to either of its branches, yielding the set of generated photo-currents

$$i_{I\pm}(t) = \frac{R_{pd}}{2}\left[P_{LO}(t) + P_{in}(t) \pm 2\sqrt{P_{in}(t)\,P_{LO}(t)}\cos\left(2\pi f_{IF} t + \Delta\theta(t)\right)\right] \tag{2.23}$$

$$i_{Q\pm}(t) = \frac{R_{pd}}{2}\left[P_{LO}(t) + P_{in}(t) \pm 2\sqrt{P_{in}(t)\,P_{LO}(t)}\sin\left(2\pi f_{IF} t + \Delta\theta(t)\right)\right] \tag{2.24}$$

where $f_{IF} = f_{in} - f_{LO}$ and $\Delta\theta(t) = \theta_{in}(t) - \theta_{LO}$ (IF stands for intermediate frequency). Notice that the 180° shifted outputs bring the benefit of DC component cancellation when these photocurrents are generated by two pairs of balanced photodiodes and fed to transimpedance amplifiers (TIA) in a subtracting configuration, such that

$$v_I(t) = K\left(i_{I+}(t) - i_{I-}(t)\right) = 2KR_{pd}\sqrt{P_{in}(t)\,P_{LO}(t)}\cos\left(2\pi f_{IF} t + \Delta\theta(t)\right) \tag{2.25}$$

$$v_Q(t) = K\left(i_{Q+}(t) - i_{Q-}(t)\right) = 2KR_{pd}\sqrt{P_{in}(t)\,P_{LO}(t)}\sin\left(2\pi f_{IF} t + \Delta\theta(t)\right) \tag{2.26}$$

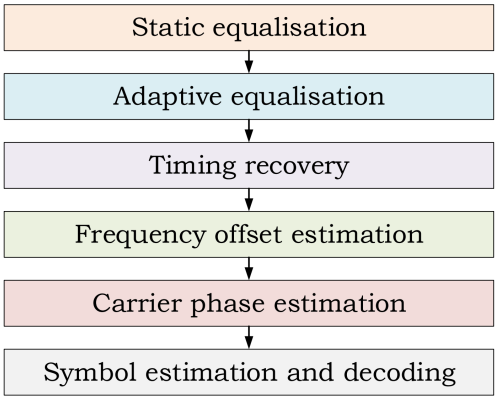


Figure 2.15 – Typical sequence of DSP stages (from top to bottom) in a coherent receiver for transmission impairments compensation

where K is the voltage-to-current gain of the TIAs (and here it is assumed to be the same for both branches, I and Q).

2.3.3 Digital signal processing algorithms

Once the optical inputs have been combined and detected, the produced signal voltages $v_I(t)$ and $v_Q(t)$ are sampled by analogue-to-digital converters (ADC), buffered and then are ready to be processed. At this point, the implementation of subsequent DSP stages has reported immense benefits in the field of impairment compensation [93], significantly increasing the system tolerance to propagated distance, accumulated dispersion, ISI, laser phase noise, channel non-linearities, etc. Figure 2.15 shows a flow diagram that illustrates the conventional (but not unique) order of DSP algorithms in a coherent receiver. From these stages, carrier phase estimation (CPE) is the one that has a major impact on the outcome of this thesis and therefore it is discussed in Section 2.4. The rest of the stages are briefly discussed in the remaining of this section.

Static equalisation

Equalisation in digital coherent receivers is commonly split in two stages: static and adaptive. In the first case, large static filters are used to compensate for linear impairments that do not change with time. This is the case for CD, which transfer

function in the frequency domain is given by [94]

$$H_{CD}(\omega, z) = \exp\left(-j\frac{\omega^2\beta_2 z}{2}\right) \tag{2.27}$$

where z is the transmitted distance and ω is angular frequency. CD compensation (CDC) is thus performed by passing the impaired signal through a filter that implements the inverse of Equation (2.27), i.e. $G(\omega) = 1/H_{CD}(\omega, z) = H_{CD}(\omega, -z)$. For this purpose, a FIR filter is employed, which basic structure is illustrated in Figure 2.16. The discrete-time output of such filter is given by the vector multiplication

$$y(n) = \mathbf{h}^T\mathbf{x} \tag{2.28}$$

where $\mathbf{x}$ is a column vector containing the last N samples of the input signal, $\mathbf{h}$ is the N-tap filter response, and $(\cdot)^T$ indicates transposition. The tap coefficients in $\mathbf{h}$ are obtained from truncating the inverse Fourier transform of $H_{CD}(\omega, -z)$ as [95]

$$[\mathbf{h}]_n = h(n) = \frac{1}{\zeta}\exp\left(-j\frac{\pi}{\zeta}\left(n - \frac{N-1}{2}\right)^2\right) \tag{2.29}$$

where $\zeta = 2\frac{\pi\beta_2 z}{T_s^2}$, T_s is the sampling period, and $N = \lfloor\zeta\rfloor$ where the operator $\lfloor\cdot\rfloor$ rounds to the nearest integer towards minus infinity.

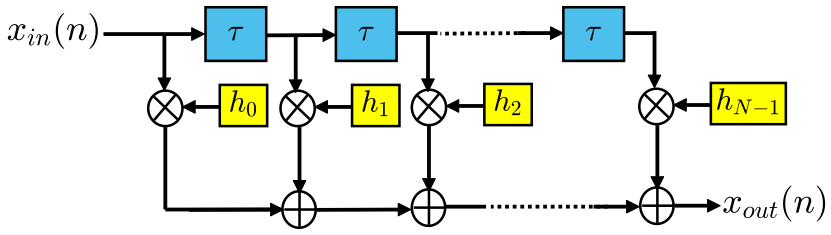


Figure 2.16 – Finite impulse response (FIR) filter structure

Adaptive equalisation

Any realistic channel exhibits a transfer function that distorts the signal and creates ISI. The receiver does not generally have an *a priori* knowledge of the

channel response and, on top of that, some of the phenomena affecting the channel are time varying, such as the fibre polarisation mode dispersion (PMD), which causes random variations of the signal's state of polarisation (SOP).

Probably the most celebrated technique for adaptive equalisation is Godard's Constant Modulus Algorithm (CMA), published in 1983 [96]. The CMA is best suited to modulation formats with symbols at a constant power level, like 4-QAM or M-ary PSK, although it is widely used as a pre-convergence algorithm with higher order QAM formats [70]. The algorithm works by iteration, first filtering the signal and then computing the updated FIR tap coefficients that minimise a cost function proportional to the deviation of the output from the ideal constant power. The algorithm has also been adapted to the dual polarisation by allowing power exchange between polarisations by means of the MIMO structure shown in Figure 2.17.

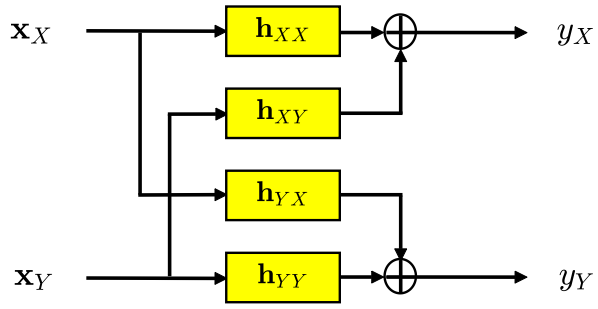


Figure 2.17 – 2x2 Multiple-input multiple-output (MIMO) filter structure for PMD compensation in polarisation-diversity coherent receivers

If the input sample vectors are denoted by $\mathbf{x}_X$ and $\mathbf{x}_Y$ for the X and Y polarisations, respectively, at the first stage the output samples are computed as

$$y_X(n) = \mathbf{h}_{XX}^H \mathbf{x}_X + \mathbf{h}_{XY}^H \mathbf{x}_Y \tag{2.30}$$

$$y_Y(n) = \mathbf{h}_{YX}^H \mathbf{x}_X + \mathbf{h}_{YY}^H \mathbf{x}_Y \tag{2.31}$$

where the subscripts in $\mathbf{h}_{out,in}$ refer to the input/output polarisations and the superscript $(\cdot)$ indicates the Hermitian or complex transpose. For constant power

formats and assuming signal power normalisation, the error terms are obtained as

$$e_X \leftarrow 1 - |y_X(n)|^2 \tag{2.32}$$

$$e_Y \leftarrow 1 - |y_Y(n)|^2 . \tag{2.33}$$

Finally, the FIR-taps are updated as follows

$$\mathbf{h}_{XX} \leftarrow \mathbf{h}_{XX} + \mu e_X \mathbf{x}_X y_X^* \tag{2.34}$$

$$\mathbf{h}_{XY} \leftarrow \mathbf{h}_{XY} + \mu e_X \mathbf{x}_Y y_X^* \tag{2.35}$$

$$\mathbf{h}_{YX} \leftarrow \mathbf{h}_{YX} + \mu e_Y \mathbf{x}_X y_Y^* \tag{2.36}$$

$$\mathbf{h}_{YY} \leftarrow \mathbf{h}_{YY} + \mu e_Y \mathbf{x}_Y y_Y^* \tag{2.37}$$

where $(\cdot)$ denotes complex conjugate and μ is the step-size parameter that determines the trade-off between stability and speed of convergence of the algorithm.

Variations of the CMA structure have been proposed for higher-order QAM formats [97,98] as well as alternative implementations in the frequency domain [99].

2.4 Carrier recovery

Once all the previous stages of normalisation, orthogonalisation, and equalisation have been performed, the resulting signal contains, ideally, impairments of angular nature only, mostly a frequency offset and a phase noise term (as long as nonlinear effects are neglected). Carrier recovery can be then split into a former stage of coarse frequency offset compensation and a posterior stage of finer carrier phase estimation (CPE). The first stage can be performed by finding the frequency f_o at which the Fourier transform of the signal exhibits a maximum, and multiplying the signal by $\exp(-j2\pi f_o t)$ with $t = kT$. In some scenarios, however, such as with strong noise or the use of probabilistic shaping, it might be necessary to apply a nonlinear function to the signal (e.g. squaring, 4-th power) prior to finding the maximum in the transform. CPE is also a critical stage, highly dependent on the modulation format, and whose goal is to track the phase variations due to the combined transmitter and LO lasers' linewidth. The different CPE algorithms fall within two categories: non-data aided (NDA) and decision-directed (DD). These

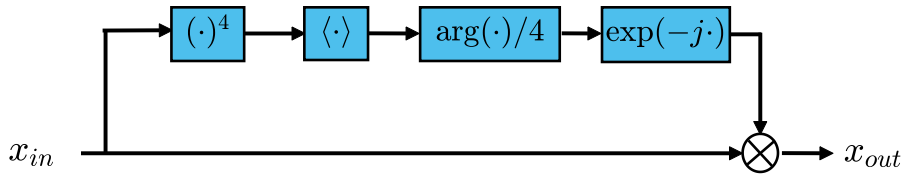


Figure 2.18 – Block diagram of the Viterbi & Viterbi phase estimator for QPSK. The $\langle\cdot\rangle$ indicates complex-valued arithmetic mean.

are discussed in Sections 2.4.1 and 2.4.2, respectively.

2.4.1 The Viterbi & Viterbi algorithm

The Viterbi & Viterbi algorithm [100] exploits the fact that given the set of received PSK symbols

$$y(t) = A\cos\left(2\pi f_c t + \frac{2\pi m}{M} + \theta_{pn}(t)\right) \tag{2.38}$$

where A is the amplitude, f_c the carrier frequency, m the symbol index, and $\theta_{pn}(t)$ the phase noise, fall onto the same point in the complex plane when they are raised to the M^{th} power. This is effectively a modulation removal, after which the symbols can be averaged out and the phase noise component extracted and removed from the original signal. Figure 2.18 illustrates the V&V algorithm for the case when $M = 4$, i.e. for a QPSK scenario.

Curiously, the V&V and Godard's CMA algorithms share some similarities: both were published in the early eighties and both were originally conceived to exploit the constant-modulus property of the M-PSK modulation format (especially QPSK). Due to their simplicity and robustness, both became the most implemented solutions for CPE and equalisation stages, respectively, and are still widely employed in modern commercial systems.

2.4.2 Decision directed CPE

As it could be noticed, the V&V algorithm did not require any knowledge of the data symbols being transmitted. On the other hand, a different family of algorithms do need to perform symbol decision to some extent in order to obtain a phase estimate. These are better suited for higher modulation formats such as M-QAM when $M > 4$.

Among the decision directed CPE algorithms (DD-CPE) it is worth naming two particular ones: the maximum *a posteriori* (MAP) and the blind phase search (BPS). The former is important because it is known to obtain the optimum solution to the CPE problem [101,102] but its computational cost is prohibitive. The latter, is popularly known as the *de facto* reference CPE algorithm for benchmarking. However, although it was published as to be "hardware efficient", in the literature it is not actually considered a suitable solution for real-time high-speed hardware implementation.

The DD-CPE problem is reviewed with more detail in Chapter 4. It will be seen that after CPE, the signal presents a residual amount of phase noise, which can be better described with the PCAWGN channel model that is, in turn, going to be presented in detail in Chapter 3.

2.4.3 Summary

This chapter has revisited some of the most important components and methods in typical optical communications systems. Starting with the transmitter side, the semiconductor laser with its rate equations and the two main modulation techniques, direct and external, have been introduced. Then the optical fibre channel is briefly described, mentioning some of the major impairments such as attenuation and dispersion. The chapter has ended with a review of optical receivers, particularly the coherent receiver followed by some popular DSP compensation algorithms.

Chapter 3

Models and approximations of the partially coherent AWGN channel

3.1 Introduction

Communication channels can be distinguished according to their degree of coherence, or in other words, their ability to retrieve the carrier signal phase. In many cases, for the receiver it is impossible, or very difficult, to obtain the phase reference of the carrier, due to issues like frequency hopping, phase noise, frequency offsets, non-coherent detection, etc. In such case, it is said to be a noncoherent channel and it has been extensively studied in the past [103–106]

A slightly different case is when the receiver has some mechanism to obtain the carrier signal phase. If this mechanism was perfect, it would render a fully coherent channel, but real systems are not perfect and the fully coherent channel has no practical use. Instead, the imperfect channel turns into a partially coherent channel, in which the signal phase is recovered but with an amount of uncertainty, that is, an angular noise contribution. Precisely, it is considered that in an optical system that performs coherent detection, the CPE stage will be corrupted by the AWGN and thus yielding an inexact phase recovery, introducing a residual phase error. In this work, polar coordinates are considered the best way to tackle the modelling of partially coherent channels, mainly because the linear and angular

noise contributions can be better split without interfering between each other, while bringing a reduction in complexity that may enable other applications that are restricted in nowadays systems.

3.2 Problem statement

Having introduced the different types of AWGN channels according to their degree of coherence, consider now the discrete-time, memoryless PCAWGN channel where, at any given instant of time $t = kT$, it is described by the expression

$$y = x \exp(j\theta) + n \tag{3.1}$$

where x and y are complex-valued random variables representing the transmitted symbol and observed sample, respectively, θ is the phase noise term, n is the linear complex additive white Gaussian noise (AWGN) with zero mean and variance N_0, that is

$$n \in \mathbb{C} \tag{3.2}$$

$$\Re(n) \sim \mathcal{N}\left(0, \frac{N_0}{2}\right) \tag{3.3}$$

$$\Im(n) \sim \mathcal{N}\left(0, \frac{N_0}{2}\right), \tag{3.4}$$

and $j = \sqrt{-1}$.

The symbols x are drawn from the discrete constellation $\mathcal{X}$ of size M, and for the cases considered in this work, unless stated otherwise, each symbol is equiprobable, i.e. the probability mass function (PMF) of x is given by

$$p(x_m) = \frac{1}{M}, \quad m \in \{0, 1, ..., M-1\} \tag{3.5}$$

The observed variable y is considered to be continuous, although in practical scenarios it would commonly be a discrete variable bounded to a limited resolution given by the number of bits of the ADC at the receiver. Thus, we define in polar

coordinates y: $(|y|, \varphi)$ with

$$|y| = \sqrt{\Re(y)^2 + \Im(y)^2} \qquad 0 \leq |y| < \infty \tag{3.6}$$

$$\varphi = \arg(y) = \tan^{-1}\left(\frac{\Im(y)}{\Re(y)}\right) \qquad -\pi \leq \varphi < \pi \tag{3.7}$$

Notice that in Equation (3.6) the domain for $|y|$ is considered to be $[0, \infty)$. While this is useful for a theoretical analysis, in practice the maximum value will be limited by an upper maximum value, $|y|_{max}$, which could be determined by the amplifiers maximum output value (i.e. clamped to the supply voltage), or by the ADC maximum allowed value, etc.

For the cases described in the following chapters, the SNR is going to be defined in dB as

$$\text{SNR} = 10 \cdot \log_{10}\left(\frac{P_S}{N_0}\right) \tag{3.8}$$

where P_S is the average transmitted power, that is

$$P_S = \mathbb{E}\left[|x|^2\right], \tag{3.9}$$

where $\mathbb{E}[\cdot]$ denotes the operator for statistical expectation.

3.3 Exact solution

There is a natural benefit that arises from modelling the channel described in Equation (3.1) in polar coordinates, and it is the fact that the phase noise contributes only to the angular dimension, leaving the modulus component unaffected. Therefore, it is convenient to divide the problem in two parts: first, analyse the impact of the AWGN on the modulus and angular distributions and, secondly, include the effects of the phase noise by applying convolution.

Case I: AWGN-only channel

In this first case, the phase noise in Equation (3.1) is neglected, therefore, the process θ is zero at all times. In this new scenario, the channel conditional probab-

ility $p(y|x)$ is well known and given by [107]

$$p(y|x) = \frac{|y|}{\pi N_0} \exp\left(-\frac{|x|^2 + |y|^2 - 2|x|\rho\cos\Delta\varphi_0}{N_0}\right) \tag{3.10}$$

where $\Delta\varphi_0$ is the modulo-π angular difference between x and y, i.e.

$$\Delta\varphi_0 = \varphi - \arg(x) \tag{3.11}$$

with $\Delta\varphi_0 \in [-\pi, \pi)$. From Equation (3.10) it is possible to extract the marginal probability distribution function (PDF) of the observed amplitude $|y|$, which is the also well-known Rician distribution

$$p(|y||x) = \frac{|y|}{N_0/2} \exp\left(-\frac{|x|^2 + |y|^2}{N_0}\right) I_0(\kappa) \tag{3.12}$$

where $I_0(\kappa)$ is the modified Bessel function of the first kind and its argument κ is defined as

$$\kappa = \frac{|x||y|}{N_0/2}. \tag{3.13}$$

Moreover, assuming that $|y|$ has been observed, it is possible to express the *a posteriori* probability (APP) of the angular difference $\Delta\varphi_0$ as follows

$$p(\Delta\varphi_0|x, |y|) = \frac{\exp[\kappa\cos(\Delta\varphi_0)]}{2\pi I_0(\kappa)} \tag{3.14}$$

which corresponds to the *von Mises* distribution $\mathcal{M}(\mu, \kappa)$ with probability density function

$$p_{\mathcal{M}}(\phi; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} e^{\kappa\cos(\phi-\mu)} \tag{3.15}$$

with mean $\mu = 0$ and a *concentration parameter* κ given as in Equation (3.13).

Case II - The PCAWGN channel

The first thing that must be noticed in order to find the joint probability distribution of the AWGN and the phase noise is that the modulus distribution is not affected by the phase noise and, therefore, its PDF $p(|y||x)$ is the same as in Equation (3.12). Secondly, a choice is required to model the phase noise process θ

in Equation (3.1). In this regard, there are mainly two options that are recurrent in most literature [105, 108, 109]. One is the wrapped Gaussian or wrapped normal distribution $\mathcal{WN}(\mu, \sigma)$ with the following PDF

$$p_{\mathcal{WN}}(\phi; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \exp\left[-\frac{(\phi - \mu + 2\pi k)^2}{2\sigma^2}\right] \tag{3.16}$$

which obviously presents the practical limitation imposed by the infinite sum over the index k (for the simulations presented in this work k is kept equal to 3). Alternatively, the truncated version of Equation (3.16) can be also considered, using the single value for $k = 0$ only, which is equivalent to the unwrapped Gaussian or normal distribution $\mathcal{N}'(\mu, \sigma)$ with PDF given by

$$p_{\mathcal{N}'}(\phi; \mu, \sigma) = \frac{K'}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\phi - \mu)^2}{2\sigma^2}\right] \tag{3.17}$$

where, for this particular case, the term $K' \neq 1$ is a normalisation factor to ensure that $\int_{-\pi}^{\pi} p_{\mathcal{N}'}(\phi; \mu, \sigma)\, d\phi = 1$.

Another popular choice to model the PDF of the phase noise θ is the von Mises distribution —also known as the Tikhonov distribution—, given by Equation (3.15). The Tikhonov distribution $\mathcal{M}(0, \kappa)$ is thus a good approximation of $\mathcal{WN}(\mu, \sigma)$, in which the relation between the circular variance σ^2 and the concentration parameter κ is given by [108]

$$\kappa \approx \frac{1}{\sigma^2} \tag{3.18}$$

Figure 3.1 shows a comparison of the wrapped Gaussian (solid black), the truncated Gaussian (dotted blue), and the Tikhonov (red dashed) distributions for several values of σ. It can be seen that for smaller amounts of phase noise ($\sigma = 0.5,\ \sigma = 1$) the truncated version still coincides perfectly with the wrapped one, while the Tikhonov shows best fit for values around $\sigma = 2$. On the other hand, for higher values of σ where the distribution approaches to uniform, the Tikhonov approximates better than the truncated Gaussian.

In what follows, the phase noise term will be considered to follow a Tikhonov distribution with zero mean and concentration parameter α, i.e. $\theta \sim \mathcal{M}(0, \alpha)$. Now the angular difference between the variables x and y results from the process

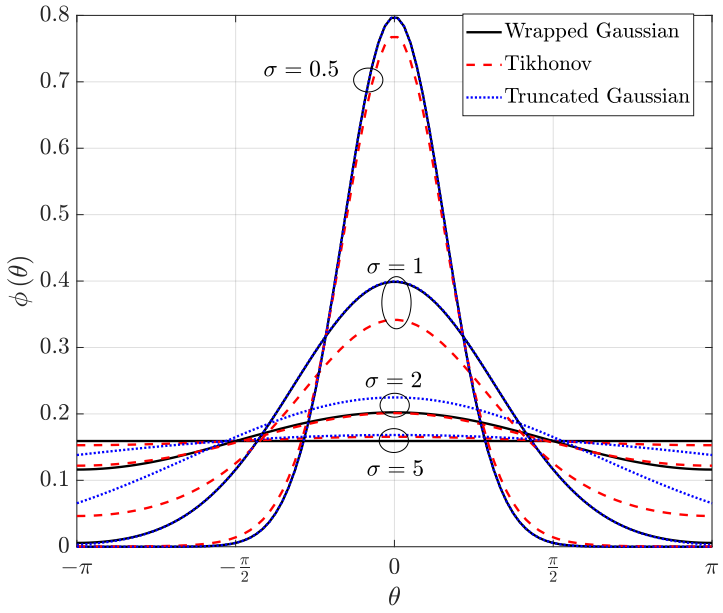


Figure 3.1 – Comparison of the wrapped Gaussian distribution (solid black) with the truncated/unwrapped (dotted blue) and the Tikhonov (dashed red) distributions for different values of σ.

addition

$$\Delta\varphi = \Delta\varphi_0 + \theta \tag{3.19}$$

and as such, its PDF $p(\Delta\varphi|x,|y|)$ is obtained from the convolution of the two generating distributions, $p(\Delta\varphi_0|x,|y|)$ and $p(\theta)$, and it is expressed as follows [110, Eq. 3.5.43]

$$p(\Delta\varphi|x,|y|) = \frac{I_0\left(\sqrt{\kappa^2+\alpha^2+2\kappa\alpha\cos(\Delta\varphi)}\right)}{2\pi I_0(\kappa)\, I_0(\alpha)}. \tag{3.20}$$

Finally, the joint PDF is obtained by multiplying Equations (3.12) and (3.20), which yields

$$p(y|x) = \frac{|y|}{\pi N_0}\exp\left(-\frac{|x|^2+|y|^2}{N_0}\right)\cdot\frac{I_0(\nu)}{I_0(\alpha)} \tag{3.21}$$

with

$$\nu = \sqrt{\kappa^2+\alpha^2+2\kappa\alpha\cos(\Delta\varphi)}. \tag{3.22}$$

For the rest of this work, Equation (3.21) will be referred to as the *exact* solution of the PCAWGN channel conditional probability. It is also noticeable that this result is similar to the one shown in [109, Eq. 3] although in the latter case the authors

took into account other factors such as the portion of power in a pilot tone sent with the signal and the one-sided loop bandwidth-sampling period product.

3.4 Approximations of the PCAWGN channel

In this section, three different ways to approximate the exact solution in Equation (3.21) are presented. The first way is simply a collection of numerical approximations for the Bessel functions involved in the computation. The second method is a previously proposed metric that will be used in simulations for the sake of comparison. The third method, is a new metric in polar coordinates that is presented as the main original contribution of this thesis.

3.4.1 Numerical approximations

Due to the presence of non-trivial modified Bessel functions, it is sometimes required to rely on numerical approximations to evaluate the model in Equation (3.21). For instance, Figure 3.2 shows the curve $\log_{10}\left[I_0\left(\kappa\right)\right]$ calculated with MATLAB, which overflows for input values around $\kappa \approx 701$ and above (region shaded in yellow). Such values are easily achieved for both arguments of the two $I_0\left(\cdot\right)$ functions in Equation (3.21), α and ν (which in turn depends on κ and α itself).

Moreover, because such parameters are inversely proportional to the noise variance, i.e. $\kappa \propto N_0^{-1}$ and $\alpha \propto \sigma^{-2}$, they are more likely to produce overflows at low values of noise where, in contrast, the function should be expected to behave as well as in the absence of noise.

The overflow issues can be overcome by using the asymptotic approximation of $I_v\left(z\right)$ such as the infinite power series (see e.g. [111, Eqs. 10.40.1 and 10.17.1])

$$I_v\left(z\right) \approx \frac{e^z}{\sqrt{2\pi z}} \sum_{k=0}^{\infty} \left(-1\right)^k \frac{a_k\left(v\right)}{z^k} \tag{3.23}$$

$$a_k\left(v\right) = \frac{\left(4v^2-1^2\right)\left(4v^2-3^2\right)\cdots\left(4v^2-\left(2k-1\right)^2\right)}{k!8^k}, \quad \text{for } k \geq 1. \tag{3.24}$$

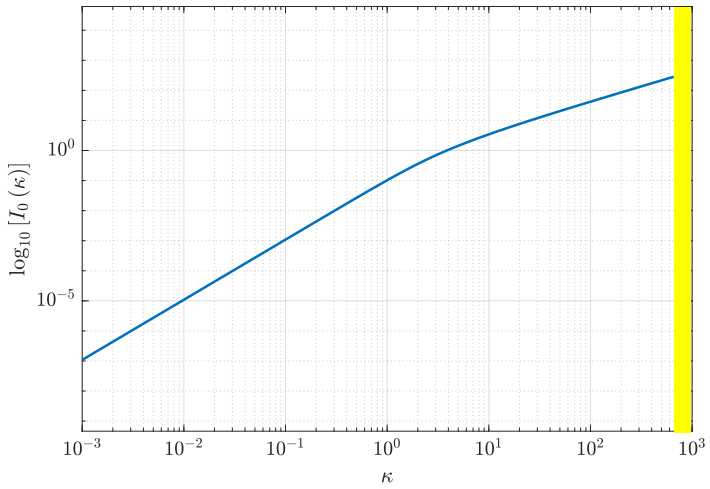


Figure 3.2 – $\log_{10}$ of the modified Bessel function of the first kind, $I_0(\kappa)$, calculated in MATLAB. The yellow-shaded area represents the range of inputs (x-axis) for which the function produces overflow (starting at $\kappa \approx 701$).

For simulation purposes, up to four expansion terms have been used to approximate Equation (3.23), yielding the following expression

$$I_0(z) \approx \hat{I}_0(z) = \frac{e^z}{\sqrt{2\pi z}}\left(1 + \frac{1}{8z} + \frac{9}{128z^2} + \frac{25}{3072z^3}\right) \tag{3.25}$$

whose error associated to the truncated expansion does not exceed the first neglected term in absolute value and has the same sign. In the case used for $k = 3$, the error is bound by

$$\delta\hat{I}_0(z) \leq \frac{e^z}{\sqrt{2\pi z}}\left(2\frac{49}{98304}|z|^{-4}\exp\left(\frac{|z|^{-4}}{4}\right)\right) \tag{3.26}$$

Additionally, in some cases there might be an overflow or indetermination caused by the quotient $I_0(\nu)/I_0(\alpha)$ in Equation (3.21). Hence, in such cases it was found convenient to resolve the issue by first applying the logarithm operator on both sides of the equation, as

$$\log(\hat{p}(y|x)) \approx \log\left(\frac{|y|}{\pi N_0}\right) - \frac{|x|^2 + |y|^2}{N_0} + \log\left(\hat{I}_0(\nu)\right) - \log\left(\hat{I}_0(\alpha)\right) \tag{3.27}$$

and then obtaining the desired result by applying the reverse operation:

$$\hat{p}\left(y|x\right) = \exp\left[\log\left(\hat{p}\left(y|x\right)\right)\right]. \tag{3.28}$$

3.4.2 Foschini's *jitter distance*

Foschini *et al.* considered the problem of designing constellations of two-dimensional symbols in the presence of AWGN combined with the so-called carrier phase *jitter* [112]. The authors formulate the problem from a vectorial point of view, where they write an equivalent to Equation (3.1) of the form

$$\mathbf{y} = R\mathbf{x} + \mathbf{n} \tag{3.29}$$

where

$$\mathbf{y} = \begin{pmatrix} y_i \\ y_q \end{pmatrix}, \ \mathbf{x} = \begin{pmatrix} x_i \\ x_q \end{pmatrix}, \ \mathbf{n} = \begin{pmatrix} n_i \\ n_q \end{pmatrix} \tag{3.30}$$

and where R is the rotation matrix defined as

$$R = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}. \tag{3.31}$$

The PDF for the channel output conditioned on the transmission of symbol $\mathbf{x}_m$ is given by

$$p\left(\mathbf{y}|\mathbf{x_m}\right) \triangleq p_m\left(\mathbf{y}\right) = \int_{-\pi}^{\pi} p_m\left(\mathbf{y}|\phi\right) p\left(\phi\right) d\phi, \tag{3.32}$$

where

$$p\left(\mathbf{y}|\phi\right) = \frac{1}{2\pi N_0} \exp\left[-\frac{1}{2N_0}\left\|\mathbf{y} - R\mathbf{x}_m\right\|^2\right] \tag{3.33}$$

$$= \frac{1}{2\pi N_0} \exp\left\{-\frac{1}{2N_0}\left[\left\|\mathbf{y} - \mathbf{x}_m\right\|^2 + 2\left\langle \mathbf{y}, \mathbf{x}_m\right\rangle - 2\left\langle \mathbf{y}, R\mathbf{x}_m\right\rangle\right]\right\} \tag{3.34}$$

where the angle brackets $\langle\cdot,\cdot\rangle$ denote the *inner* vector product. Finally, using the Tikhonov distribution of Equation (3.14) in Equation (3.32), under the assumptions of sufficiently large α and sufficiently small N_0, and after some algebraic manipula-

tions they reach the result (and now back to the complex-valued variables form)

$$p\left(y|x\right) \approx \frac{1}{2\pi N_0} \exp\left[-\frac{1}{2N_0} d^2\left(y,x\right)\right] \tag{3.35}$$

where the *effective* squared distance d^2 is obtained as

$$d^2\left(y,x\right) \triangleq |y-x|^2 + 2\left\langle y,x\right\rangle + 2\alpha N_0 - 2\left|\sqrt{|y|^2|x|^2 + 2\alpha N_0\left\langle y,x\right\rangle + \left(\alpha N_0\right)^2}\right|, \tag{3.36}$$

which was coined by the authors as the *jitter* distance. However, in what follows, the result shown in Equations (3.35) and (3.36) will be referred to as the Foschini's metric, model, or approximation.

3.4.3 A proposed approximation in polar coordinates

Differently to what has been described while developing Foschini's metric, in this thesis it is proposed an alternative approximation of the PCAWGN channel that exploits the properties of polar coordinates. In this section, the novel metric will be described and in the rest of the chapter it will be evaluated along with both the exact solution and the Foschini's metric, aiming to expose a fair comparison.

Starting with the amplitude component, for high values of SNR the Rician distribution in Equation (3.12) approaches a Gaussian distribution $\sim \mathcal{N}\left(|x|, \frac{N_0}{2}\right)$ with PDF given by

$$p\left(|y||x|\right) \approx \frac{K_\rho}{\sqrt{\pi N_0}} \exp\left(-\frac{\left(|y|-|x|\right)^2}{N_0}\right) \tag{3.37}$$

where the constant K_ρ has been introduced to ensure that the total probability within $|y| \in [0,\infty)$ is unity. In a similar way, the phase distribution described in Equation (3.20) as the convolution of two Tikhonov distributions can be approximated by another Tikhonov or by a wrapped Gaussian ([110, p. 44]) which, in turn, is here approximated by truncation to the following unwrapped Gaussian with PDF

$$p\left(\Delta\varphi|x,|y|\right) \approx \frac{K_\varphi}{\sigma_\Phi\sqrt{2\pi}} \exp\left(-\frac{\Delta\varphi^2}{2\sigma_\Phi^2}\right) \tag{3.38}$$

where, again, a constant K_φ is meant to ensure that the total probability in $\Delta\varphi \in$

$(-\pi, \pi)$ is equal to 1, and where σ_Φ^2 is the variance of the variable $\Delta\varphi$. Since $\Delta\varphi$ is obtained from convolving two distributions approximated as Gaussian, according to the *central limit theorem* its variance will be given by the sum of the two variances of the processes being convolved, that is

$$\sigma_\Phi^2 = \kappa^{-1} + \alpha^{-1} = \frac{N_0}{2\,|x|\,|y|} + \sigma_p^2 \tag{3.39}$$

where it has been assumed that for $\sigma_p^2 \ll 1 \rightarrow \sigma_p^2 \approx \frac{1}{\alpha}$.

An intuitive interpretation of the first term in the RHS of Equation (3.39) is that, according to the proposed approximation, half of the power of the AWGN term will produce a Gaussian-distributed angular displacement which variance is proportionally smaller as the point of the constellation is further from the origin (i.e. inversely proportional to the product $|x|\,|y|$).

Lastly, the proposed approximated model is obtained from multiplying Equations (3.37) and (3.38), yielding

$$p\,(y|x) \approx \frac{K\gamma}{\pi N_0} \exp\left(-\frac{\Delta\rho^2 + (\gamma\Delta\varphi)^2}{N_0}\right) \tag{3.40}$$

with

$$\Delta\rho = |y| - |x| \tag{3.41}$$

$$\gamma = \left(\frac{1}{|x|\,|y|} + \frac{2\sigma_p^2}{N_0}\right)^{-\frac{1}{2}} \tag{3.42}$$

and with $K = K_\rho K_\varphi$ which, as it will be seen, for high SNR the valid assumption of $K \approx 1$ applies.

As with Foschini's metric (see Equation (3.35)), it is also remarkable the way in which Equation (3.40) resembles an Euclidean distance. Similarly to what happens in the former model, the angles are "compressed" by the effect of the factor γ such that points that lay at a longer angular distance actually appear to be equidistant.

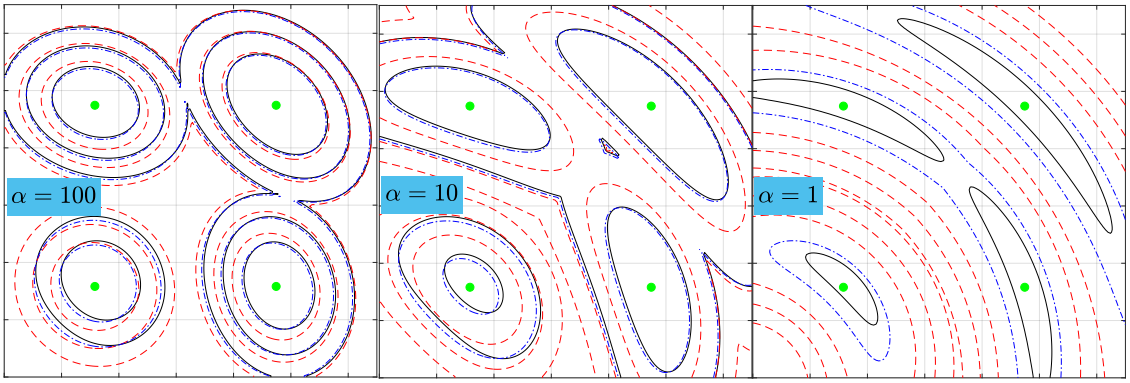


Figure 3.3 – Example of the contours for equal probability density values (also referred to as equiprobable curves) from points in the quadrant I of a 16-QAM constellation. The SNR is set to 15 dB and α takes values 100, 10, and 1. The curves are obtained with the exact PCAWGN model (solid), Foschini's metric (dashed), and the proposed model (dash-dotted).

3.5 Performance evaluation and comparison

Probability density curves

This first evaluation aims to compare the two approximations to the exact solution in terms of the contour lines of equal probability (or *equiprobable* curves). Figure 3.3 shows an example of the equiprobable curves for three values of probability density (5, 2.5, and 1) and for three different values of α (100, 10, and 1). Both the Foschini's metric and the proposed model exhibit the same shape as the exact solution, but only the latter closely matches the probability density values for most values of α. This explains why, while both approximations are good as maximum-likelihood hard-decision (ML-HD) detectors (see Figure 3.5), the proposed model yields a better calculation of channel mutual information (MI) (exposed in Section 3.5.1). The PDF for the phase noise $p(\varphi)$ with the used values of $\alpha = 100$, 10, and 1, is also shown for reference in Figure 3.4.

Hardware complexity

In modern systems that tend to implement as many DSP stages as possible, it is of crucial importance the hardware-complexity reduction that one solution might

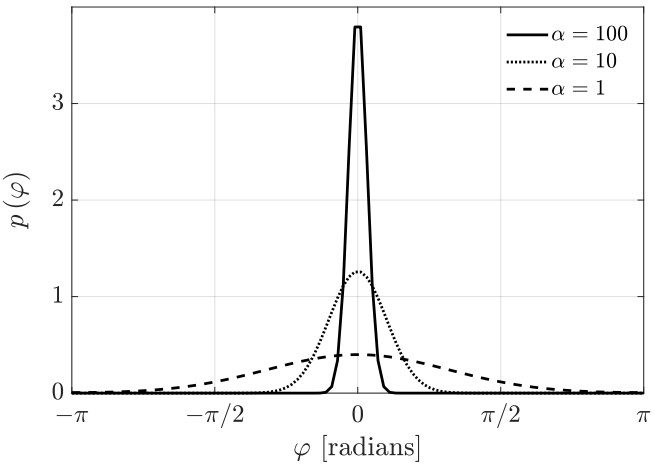


Figure 3.4 – Example of PDF for the phase noise with values of α 100, 10, and 1

bring over an existing one. In this regard, the exact solution in Equation (3.21) has the main disadvantage of the quotient with the two modified Bessel functions. These functions are no problem for desktop computers but they might result in a more expensive implementation in real-time systems based on *field-programmable gate arrays* (FPGA) or *application-specific integrated circuit* (ASIC). Even though they can be stored in a *look-up table* (LUT), the issues with the indeterminations produced by divisions of the type ∞/∞, $0/0$, etc, may persist. While both approximations, Foschini's and proposed, solve this problem by removing completely the modified Bessel functions, the main difference between them is the use of rectangular versus polar coordinates, respectively. Foschini's metric is derived by means of vector algebra operations and thus relies on inner products, requiring a total of 12 multiplications per processed symbol compared to the only 8 multiplications required by the proposed model. The number of hardware operations for the two approximations and the exact solutions is detailed in Table 3.1.

It could be reasonably argued that the hidden trick behind using polar coordinates is that their conversion from Cartesian requires some extra multiplications that here have not been accounted for. However, the celebrated Volder's algorithm (also known as CORDIC, for COordinate Rotation DIgital Computer), can be employed [113, 114]. CORDIC is a collection of iterative algorithms that implement trigonometrical functions (among other types) in a digital processor, by only using inexpensive shift-add operations. Some studies have proposed FPGA implement-

Table 3.1 – Complexity comparison of the PCAWGN channel models in terms of number of operations required per computed symbol

Model	$+$	$\times$	$\surd$	exp	$>>^*$	$I_p(z)$
Exact	2	7	1	1	1	2
Foschini	9	12	1	1	2	0
Proposed	3	8	1	1	0	0

* bit shift-register

ations [115, 116], amongst which a multiplier-free, CORDIC-based polar-domain, carrier phase tracking algorithm was experimentally tested and reported in [117].

Hard-decision symbol detection

As an initial evaluation test, the exact model and its approximations are compared by their ability to perform ML-HD symbol detection. Therefore, the figure of merit for this test is the symbol error ratio (SER) when a given set of data symbols drawn from a QAM constellation are subject to both AWGN and phase noise. The results are shown in Figure 3.5 after a Monte-Carlo simulation of 10^6 symbols for each data point, randomly generated according to Equation (3.1) with complex Gaussian N and wrapped Gaussian θ.

It is observed that for all values of α and for each of the QAM orders (4, 8, 16, and 32), both the Foschini's metric (cross markers) and the proposed approximation (circle markers) show an excellent agreement with the optimal detector (solid line), that is, the one obtained with the exact solution; in addition, the slicing function based on the Euclidean distance is also shown for reference (dashed line).

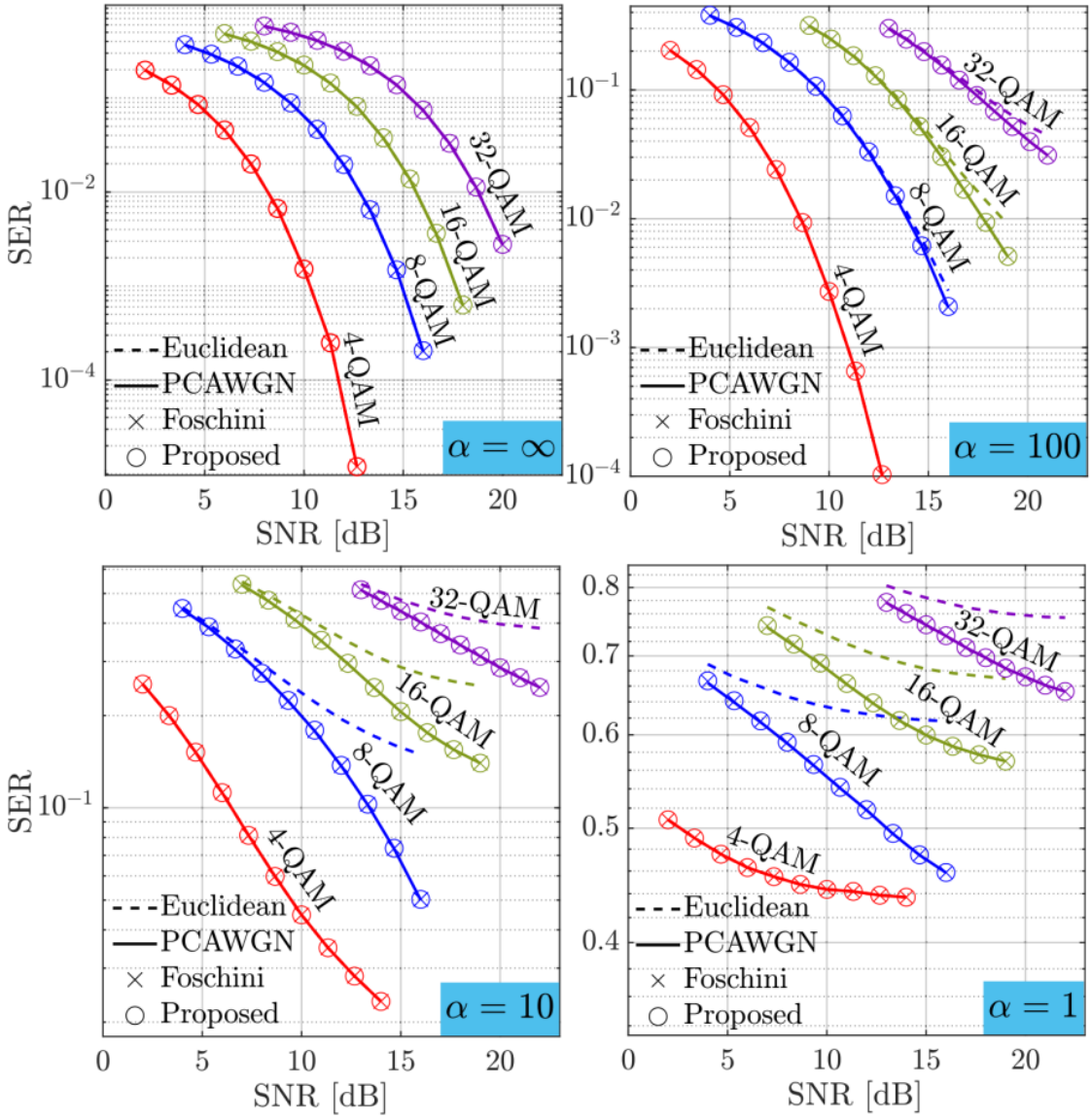


Figure 3.5 – Hard-decision symbol error rate (HD-SER) comparison using the models as maximum-likelihood detectors (ML). Results shown for the Euclidean detector (dashed), the exact model (solid line), Foschini's distance (cross markers), and the proposed metric (circle markers). α takes the values (from left to right): ∞, 100, 10, and 1. The results are shown for four QAM constellations of order 4, 8, 16, and 32. Each data point obtained from the simulation of 10^6 symbols.

3.5.1 Channel mutual information

Although hard-decision SER and BER have been widely adopted as a valid measure of a system's performance, some recent studies argue that a more sophisticated figure like the channel mutual information (MI) is a better candidate to faithfully indicate the potential performance of a given system [118]. Henceforth, in this second set of simulations, it is evaluated the amount of information loss ΔI, in bits per symbol, between the exact (I_{opt}) and the approximated (or auxiliary, I_{aux}) channel models, that is

$$\Delta I = I_{opt} - I_{aux} \tag{3.43}$$

where, for each model, I is the *mutual information* (MI) defined as the channel *achievable information rate* (AIR) and it is obtained as [119, Eq. 6]

$$I = m + \frac{1}{M}\sum_{i=1}^{M}\iint_{|y|,\varphi} p\left(y|x_i\right)\log_2\frac{q\left(y|x_i\right)}{\sum_{j=1}^{M} q\left(y|x_i\right)}d\left|y\right|\,d\varphi \tag{3.44}$$

where M is the size of the constellation set $\mathcal{X}$, $m = \log_2(M)$, and where it has been introduced the distinction between $p\left(y|x\right)$ versus $q\left(y|x\right)$ as in [120, Eq. 35] to enable the computation of both the optimal and the auxiliary-channel lower bounds. Hence, whilst $p\left(y|x\right)$ is obtained as in Equation (3.21) in both cases, $q\left(y|x\right)$ is obtained with Equation (3.21) for the optimal case, and with Equations (3.35) and (3.40) for the Foschini's and proposed channel approximations, respectively. Additionally, for the computation of MI it was necessary to apply the following normalisation to the approximated models

$$\overline{q\left(y|x_i\right)} = \frac{q\left(y|x_i\right)}{\int_{\mathcal{Y}} q\left(y|x_i\right)dy},\ \forall i \in \{1,\ldots,M\} \tag{3.45}$$

Figures 3.6 and 3.7 show the calculated ΔI curves for the α values ∞, 100, 10, and 1, respectively. The tested QAM constellations are, again, of order 4, 8, 16, and 32. In the case of AWGN-only ($\alpha = \infty$, left plot in Figure 3.6) both Foschini's metric (cross markers) and the proposed model (circle markers) exhibit identical behaviour, with the expected high penalties at low SNR and proportional to the constellation size, although rapidly vanishing as the SNR increases. Opposed to that, in the presence of phase noise the proposed model exhibits a much better performance

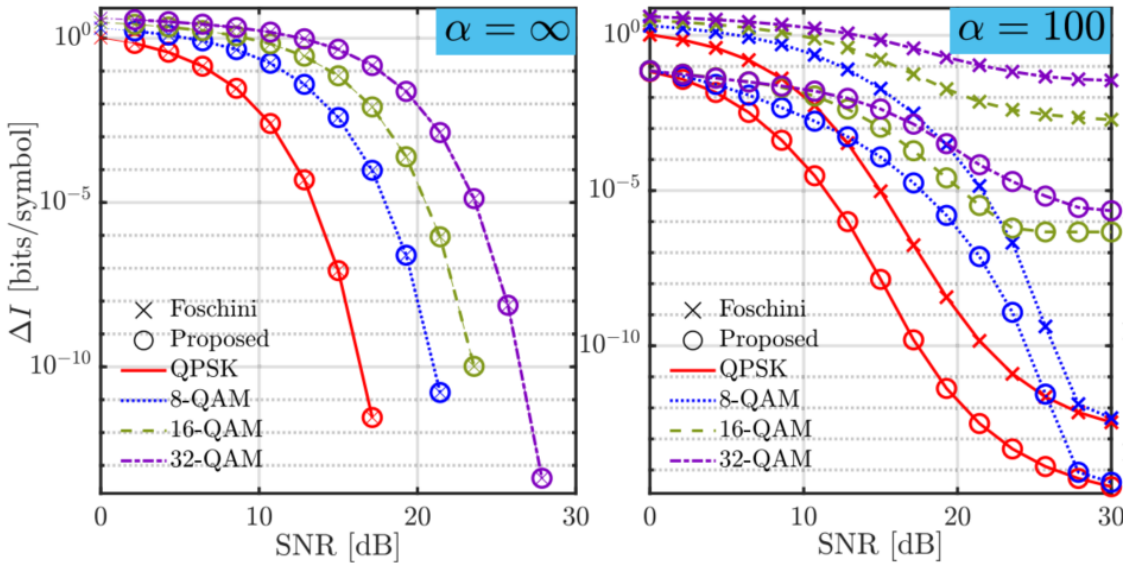


Figure 3.6 – Channel information loss (IL, ΔI) in bits/symbol for the Foschini´s metric (crosses) and the proposed model (circles) with respect to the exact solution. Results shown for QAM constellations of order 4 (solid line), 8 (dotted line), 16 (dashed line), and 32 (dash-dotted line). Two values of α tested: ∞ (left, AWGN only) and 100 (right).

(lower IL) compared to the Foschini's metric. For $\alpha = 100$ (Figure 3.6, right) the improvement is particularly significant at lower values of SNR. For instance, while the proposed metric presents an improvement of 0.25 bits/symbol at 0 dB of SNR for QPSK, a 1 bit/symbol improvement is shown at 5.5 dB, 10 dB, and 14 dB of SNR for 8-QAM, 16-QAM, and 32-QAM, respectively.

3.6 Conclusions

After the results obtained in these simulations, it can be concluded that the proposed model of the PCAWGN channel not only reduces the hardware-complexity for both the exact solution and the previously published Foschini's metric, but it also outperforms the latter in terms of its ability to compute the channel MI.

Consequently, the proposed model is envisaged as a suitable candidate for a next generation of DSP-supported optical transmitters and receivers, enabling a closer approach to the optimal implementation of densely-spaced coherent channels

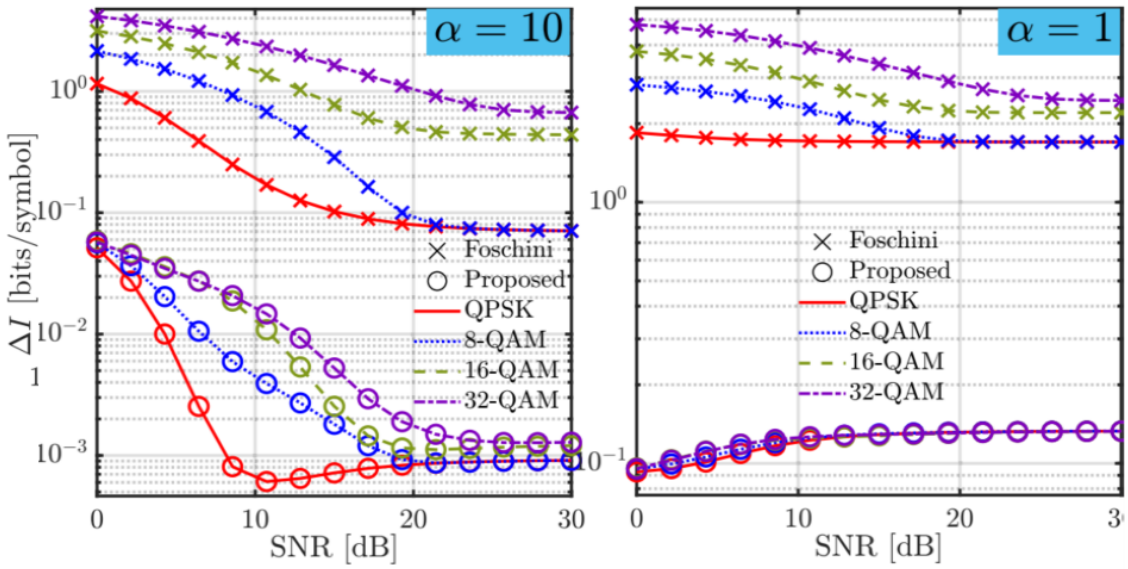


Figure 3.7 – Channel information loss (IL, ΔI) in bits/symbol for the Foschini´s metric (crosses) and the proposed model (circles) with respect to the exact solution. Results shown for QAM constellations of order 4 (solid line), 8 (dotted line), 16 (dashed line), and 32 (dash-dotted line). Two values of α tested: 10 (left) and 1 (right).

influenced by an increasing number of phase noise sources. To reinforce this belief, the next chapter is dedicated to explore and identify some of the end-to-end applications where the proposed model can be reliably used to obtain performance benefits.

Chapter 4

Applications of the approximated model

4.1 As a hard-decision slicing metric

One application for a channel conditional probability expression is the derivation of a slicing function or metric in order to perform hard symbol-decisions (HD). The slicing operation is responsible for making the symbol decisions $\hat{x}$ given the observed sample y, where the criterion is based on maximising the channel conditional probability $p(y|x)$, that is

$$\hat{x} = \arg\max_{x \in \mathcal{X}} p(y|x) \tag{4.1}$$

However, the slicing metrics can be further simplified if it is taken into account that, for the purpose of choosing one symbol against another, the actual probability value $p(y|x)$ does not matter as much as it does the relative value with respect to the compared symbol, e.g. if $p(y|x_i) > p(y|x_j)$ with $i \neq j$ and $x_i, x_j \in \mathcal{X}$, or otherwise. Therefore, probability expressions are usually reduced to their log-probability form $(-\log(P))$ where only the argument of the exponential is compared and aimed to be minimised. This argument corresponds to the actual metric of interest, the probabilistic distance or, simply, distance. For instance, the simplest case is the one for the AWGN channel, in which the argument to the exponential in the Gaussian PDF is simply the Euclidean distance squared $|y - x|^2$ divided by the noise power, N_0. Naturally, when comparing between two different symbols, the common term N_0 can be safely removed, yielding a symbol decision based on the shortest Euclidean

distance. Figure 4.1 shows an example of the symbol spread probabilities (left) for a 16-QAM constellation subject to AWGN only, and the optimum decision boundaries (right, also known as the Voronoi map). An obvious advantage that presents the Euclidean distance is that decisions boundaries are straight lines and, hence, symbol decisions can be made by solely checking if the received sample coordinates are within predefined thresholds, only requiring inexpensive comparison hardware operations.

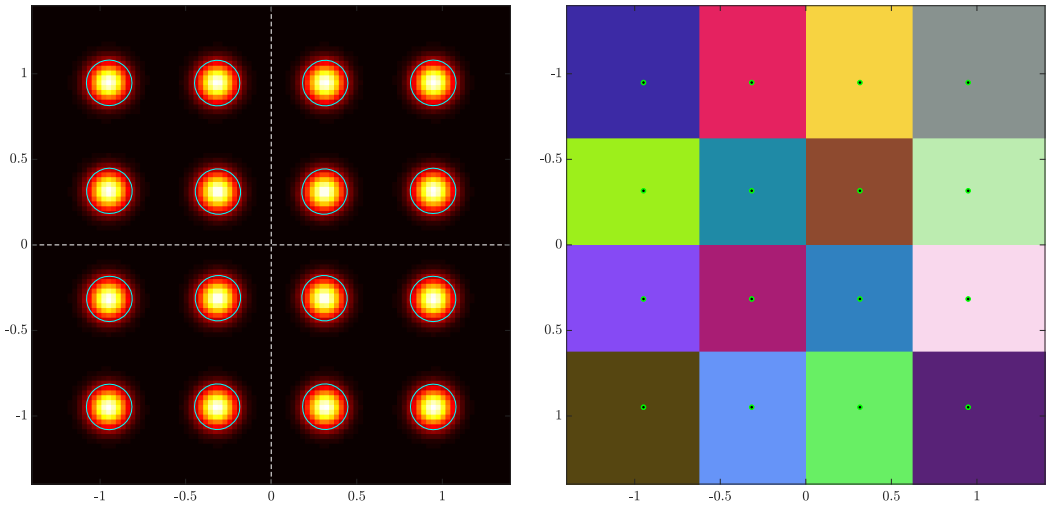


Figure 4.1 – Example of 16-QAM symbol probabilities (left) and slicing map (Voronoi map) obtained using the Euclidean distance (right) in the presence of AWGN only

On the other hand, for the PCAWGN channel, symbol decisions can be made by minimising the $-\log(P)$ function of the maximum *a posteriori* (MAP) probability, that is

$$D_{\text{MAP}} = -\log(P) = \frac{|y(k) - \hat{x}(k)\exp[j\hat{\varphi}(k)]|^2}{N_0} + \frac{(\hat{\varphi}(k))^2}{2\sigma_p^2} \tag{4.2}$$

where $\hat{\varphi}$ are phase noise estimates of the process $\varphi \sim \mathcal{WN}\left(0, \sigma_p^2\right)$ and where it has been coined as D_{MAP} because it represents the probabilistic distance of the PCAWGN channel. Unfortunately, making symbol decisions with Equation (4.2) requires a joint minimisation of $\hat{x}$ and $\hat{\varphi}$, that is

$$(\hat{x}, \hat{\varphi}) = \arg\min_{x,\varphi}(-\log(P)) \tag{4.3}$$

which indeed requires a double iteration and makes its implementation impractical for real-time applications. However, the approximated model of the PCAWGN channel presented in Chapter 3 can be of help by removing the double iteration, and it will be used to derive a reduced-complexity slicing function. Therefore, and for the sake of completeness, it is useful here to recall the proposed approximated PCAWGN channel conditional probability, given by (see Equation (3.40))

$$p_{\text{prop}}(y|x) = \frac{\gamma}{\pi N_0} \exp\left(-\frac{\Delta\rho^2 + (\gamma\Delta\varphi)^2}{N_0}\right) \tag{4.4}$$

where it must be noticed that the integration factor K has been assumed to be equal to unity. Taking the argument of the exponential of Equation (4.4) and removing the constant denominator $-N_0$ yields the proposed squared distance

$$D^2_{\text{prop}} = \Delta\rho^2 + \gamma^2\Delta\varphi^2 \tag{4.5}$$

where, while the value of γ was given in form of a squared root in Equation (3.42), now appears squared and then the distance can be expressed in a expanded form, as

$$D^2_{\text{prop}} = \Delta\rho^2 + \left(2\frac{\sigma_p^2}{N_0} + \frac{1}{|x|\,|y|}\right)^{-1} \Delta\varphi^2 \tag{4.6}$$

It can be observed that this computation is very similar to that of an Euclidean distance, where the distance squared is equal to the sum of the squared differences in the respective orthogonal coordinates (e.g. Δx and Δy). In this case, however, the angular distance component $\Delta\varphi$ is weighted to reflect the fact that it has two contributions, one due to the AWGN and another one due to the phase noise.

Figure 4.2 shows a comparison between the two metrics D_{MAP} (solid black line) and the proposed D_{prop} (dashed red line), obtained by Equations (4.2) and (4.6), respectively. The plot shows contours of equidistant curves for constellations 8-QAM (top row, with SNR = 15 dB) and 16-QAM (bottom row, with SNR = 20 dB), and for three values of σ_p^2: 0, 0.01, and 0.05 in the first case, and 0, 0.005, and 0.01 in the second case. It can be observed that despite the simpler computation required, the proposed metric matches perfectly with the optimum one for the different constellations and values of phase noise variance tested.

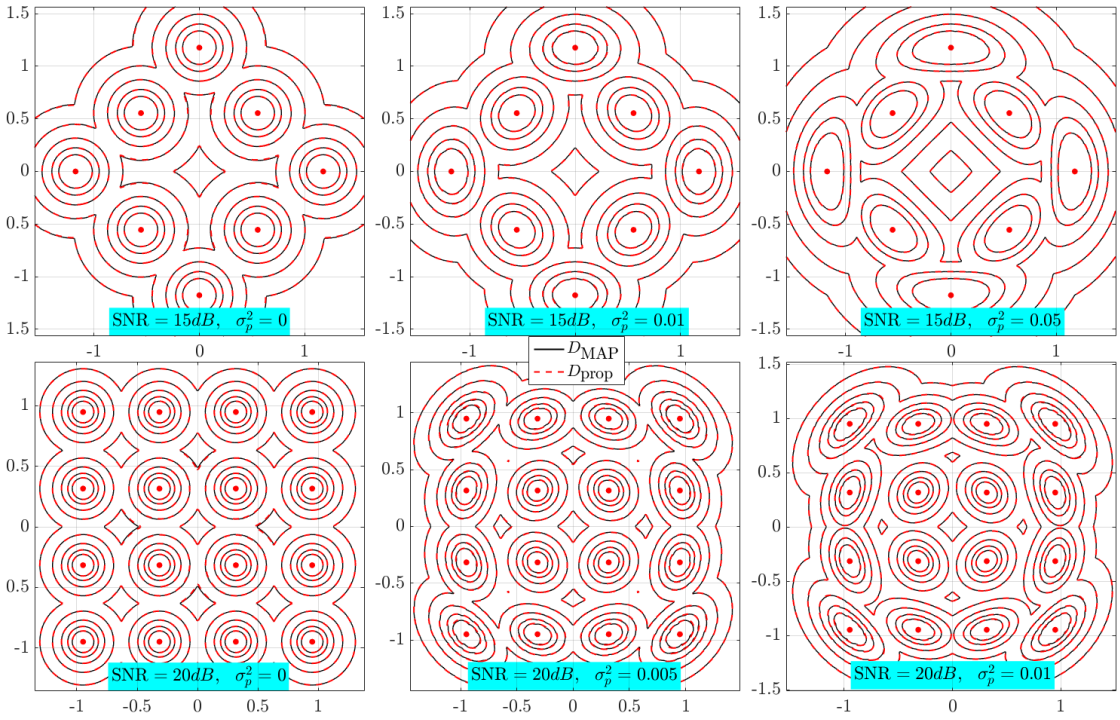


Figure 4.2 – Example of equidistant curves for 8-QAM (top) and 16-QAM (bottom) constellations computed with the optimum metric D_{MAP} (solid black line) and the proposed D_{prop} (dashed red line). In the case of 8-QAM the SNR is set to 15 dB and $\sigma_p^2 = 0, 0.01, 0.05$ whereas for the 16-QAM case the SNR is set to 20 dB and $\sigma_p^2 = 0, 0.005, 0.01$

In fact, what matters the most about a slicing metric is where the boundary decisions are located. In this regard, it can be noticed from Equation (4.6) that the calculated distance, and hence the boundary decisions, do not depend on particular values of N_0 or σ_p^2, but in the ratio between them, σ_p^2/N_0. Therefore, it makes sense to define

$$\eta = 2\frac{\sigma_p^2}{N_0}, \tag{4.7}$$

that is, the so-called *rotation factor*. Thus, any particular value of η will define a unique shape of the boundary decisions. For instance, Figure 4.3 shows the boundary decisions for a 16-QAM with three different value pairs of SNR and σ_p^2. The boundaries are the same because the rotation factor is constant among the three cases.

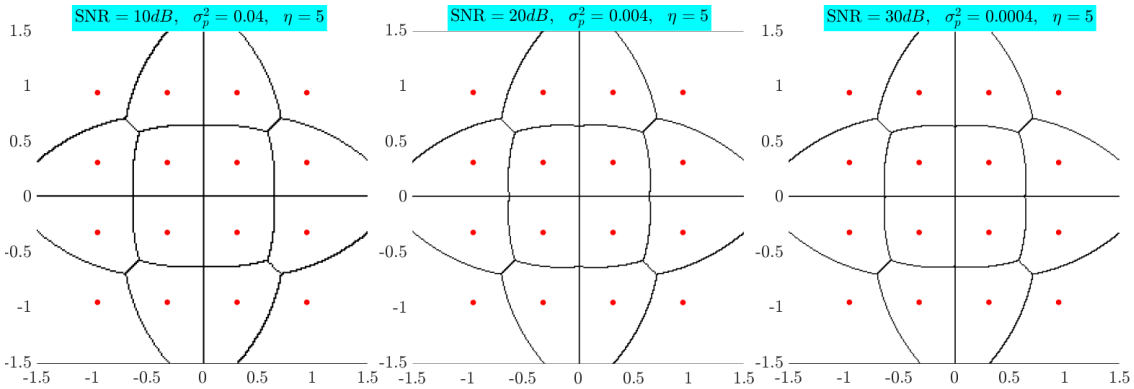


Figure 4.3 – Example of identical boundaries calculated for a 16-QAM constellation, with different values of SNR and σ_p^2 but keeping constant the rotation factor η.

4.1.1 Reducing the complexity of the proposed metric

It is possible to reduce the complexity of D_{prop} by one hardware multiplication without compromising much the performance. Consider the second term in γ which is given by $\frac{1}{|x||y|}$ and which refers to the proportion by which the variance of additive noise tangential to the point x is converted into variance of angular noise. Here there are two concepts: first, a differential linear displacement tangential to a point with radius r, say dt, will approximate to an angular rotation $d\theta \approx dt \cdot r$. Second, if dt is a normal process with variance σ_t^2, $d\theta$ will have a variance $\sigma_\theta^2 = \sigma_t^2 \cdot r^2$. Thus, it can be concluded that the linear variance is converted to an angular variance by multiplying it by an *effective* radius squared, r_{eff}^2. There can be distinguished three important cases: (i) a purely *a priori* approach, where the observed sample y is not taken into account and $r_{eff} = |x|$ [121], (ii) a non-data aided approach where the symbol x is ignored and $r_{eff} = |y|$ [122], and (iii) the actual *a posteriori* approach where the geometrical mean is used, that is $r_{eff} = \sqrt{|x|\,|y|}$ [123]. Notice that using approach (i) can reduce one multiplication compared to (ii) and (iii) if the M values for $|x_m|^2$ are stored in memory. Hence, yet another reduced-complexity version of D_{prop} is obtained by using the *a priori* version of the effective radius, reading as

$$D_{\text{prop,xx}}^2 = \Delta\rho^2 + \left(\psi + |x|^{-2}\right)^{-1} \Delta\varphi^2 \tag{4.8}$$

The same comparison shown earlier in Figure 4.2 is now performed in order to compare the two metrics D_{prop}^2 and $D_{\text{prop,xx}}^2$. The results are shown in Figure 4.4

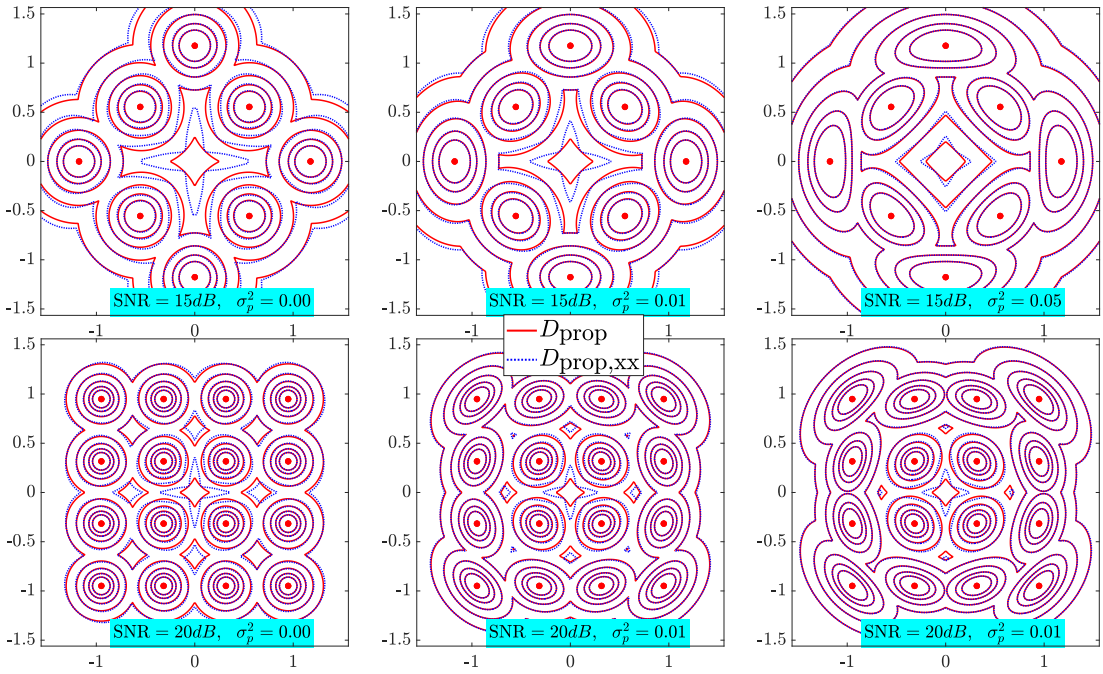


Figure 4.4 – Example of equidistant curves for 8-QAM (top) and 16-QAM (bottom) constellations computed with the proposed metric D^2_{prop} (solid blue line) and the reduced version $D^2_{\text{prop,xx}}$ (dashed red line). In the case of 8-QAM the SNR is set to 15 dB and $\sigma_p^2 = 0, 0.01, 0.05$, whereas for the 16-QAM case the SNR is set to 20 dB and $\sigma_p^2 = 0, 0.005, 0.01$

and it can be seen that there does not seem to be a significant mismatch between equidistant curves, particularly not for points further from the origin and for higher values of phase noise. However, to better quantify the impact of such mismatch, Figure 4.5 shows the results of an additional simulation, where the top plot shows the computed SER versus total SNR for D_{prop} (red) and for $D_{\text{prop,xx}}$ (blue), for 8-QAM, 16-QAM, and 64-QAM constellations. The specified total SNR in dB is computed as

$$\text{SNR}_{\text{tot}} = 10 \cdot \log_{10}\left(\frac{1}{N_0 + \sigma_p^2}\right) \tag{4.9}$$

while the three values of η, 10, 1, and 0.1, are obtained as in Equation (4.7). The plots at the bottom show the corresponding SNR_{tot} penalty in dB versus the targeted SER for all the test cases of constellation and rotation factor. It can be seen that while the mismatch is more significant for 8-QAM, with penalties of up

to 0.5 dB for SER = 10^{-4}, the penalties decrease to below 0.2 dB for 16-QAM and below 0.05 dB for 64-QAM for all targeted SER equal and above 10^{-4}.

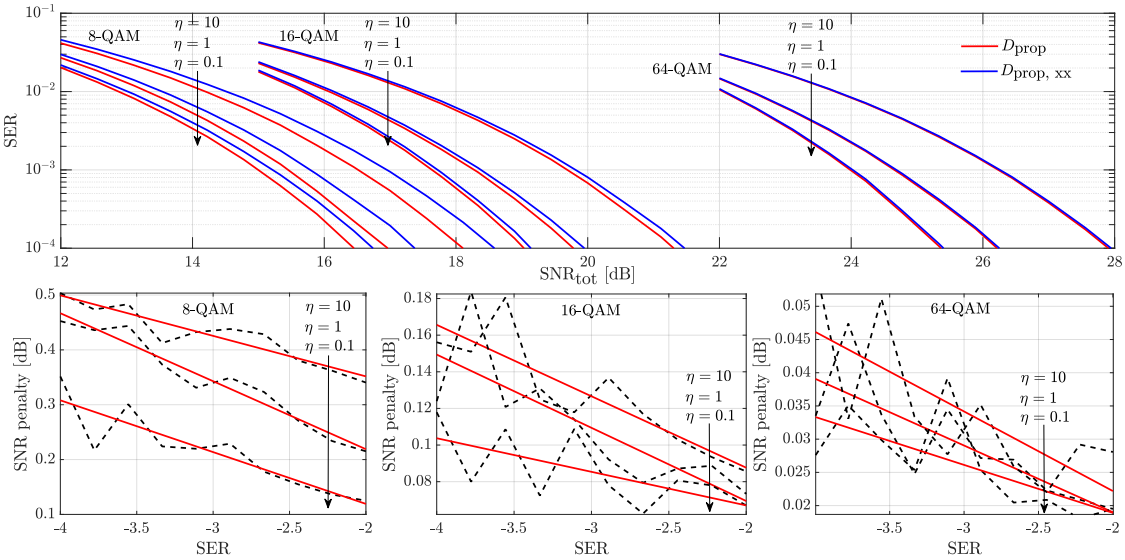
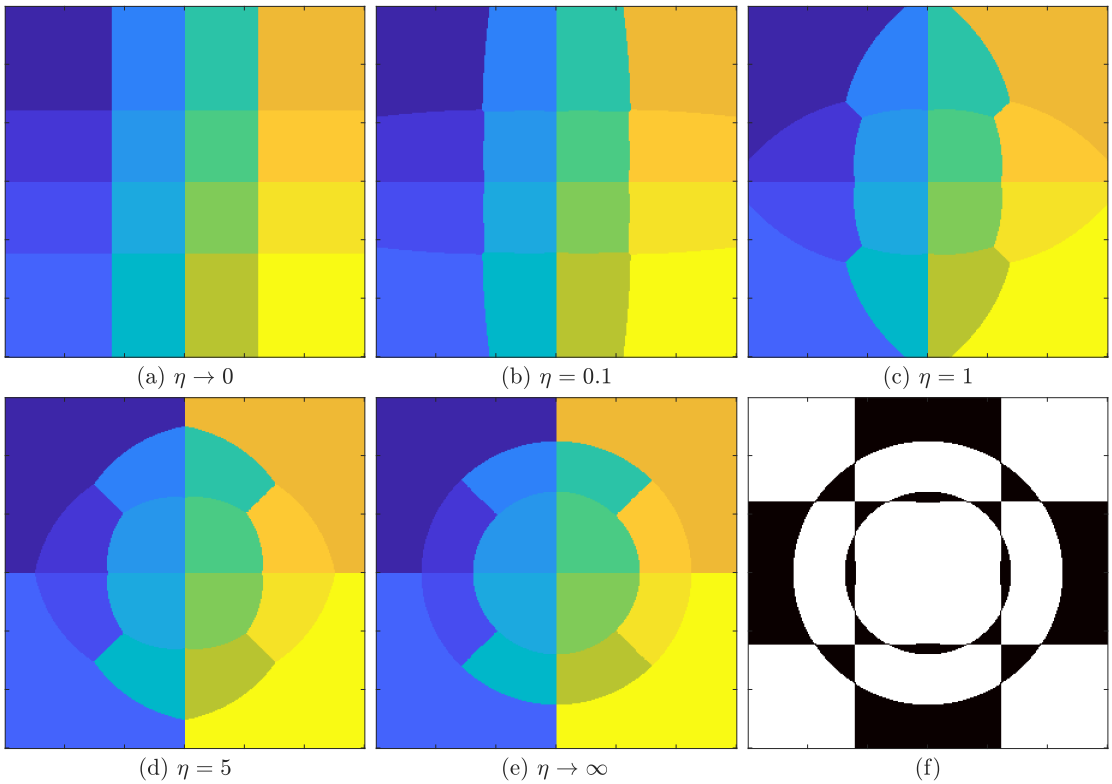


Figure 4.5 – Top: SER versus SNR_{tot} for 8-, 16-, and 64-QAM with η =10, 1, and 0.1, computed with metrics D_{prop} (red) and $D^2_{\text{prop,xx}}$ (blue). Bottom: penalty in dB for all the test cases versus targeted SER (black dashes) with linear fitting (solid red)

Another aspect that can be look at in terms of complexity is regarding the proportion of symbols for which the distance needs to be computed. Earlier in the text it was stated that a pure AWGN channel benefits from the fact that Euclidean distance generates straight lines for squared QAM constellations and, thus, slicing can be performed by comparing against predefined thresholds. For the PCAWGN, two special cases can be considered: (i) when $\eta \to 0$ ($N_0 \gg \sigma_p^2$, equivalent to AWGN only) where the boundaries are as described by the Euclidean distance (see Figure 4.6 (a)), and (i) when $\eta \to \infty$ ($N_0 \ll \sigma_p^2$), where decisions can be made entirely by comparisons in polar coordinates, i.e. by predefined angular and radial boundaries (see Figure 4.6 (e)). For intermediate values of η the boundaries do not result that straightforward, see e.g. Figure 4.6 (b, c, d). There are, however, points in the sampling space that return the same sliced symbols $\lfloor \hat{x} \rfloor$ for any value of η. These so-called *common regions* can be obtained as the coordinates where the condition $\lfloor \hat{x}_{\eta \to 0} \rfloor = \lfloor \hat{x}_{\eta \to \infty} \rfloor$ is true. Figure 4.6 (f) shows the logical space for a 16-QAM constellation with the common regions represented in white colour, and they are where symbol decisions can be made only with simple comparison operations.



(a) $\eta \to 0$ (b) $\eta = 0.1$ (c) $\eta = 1$

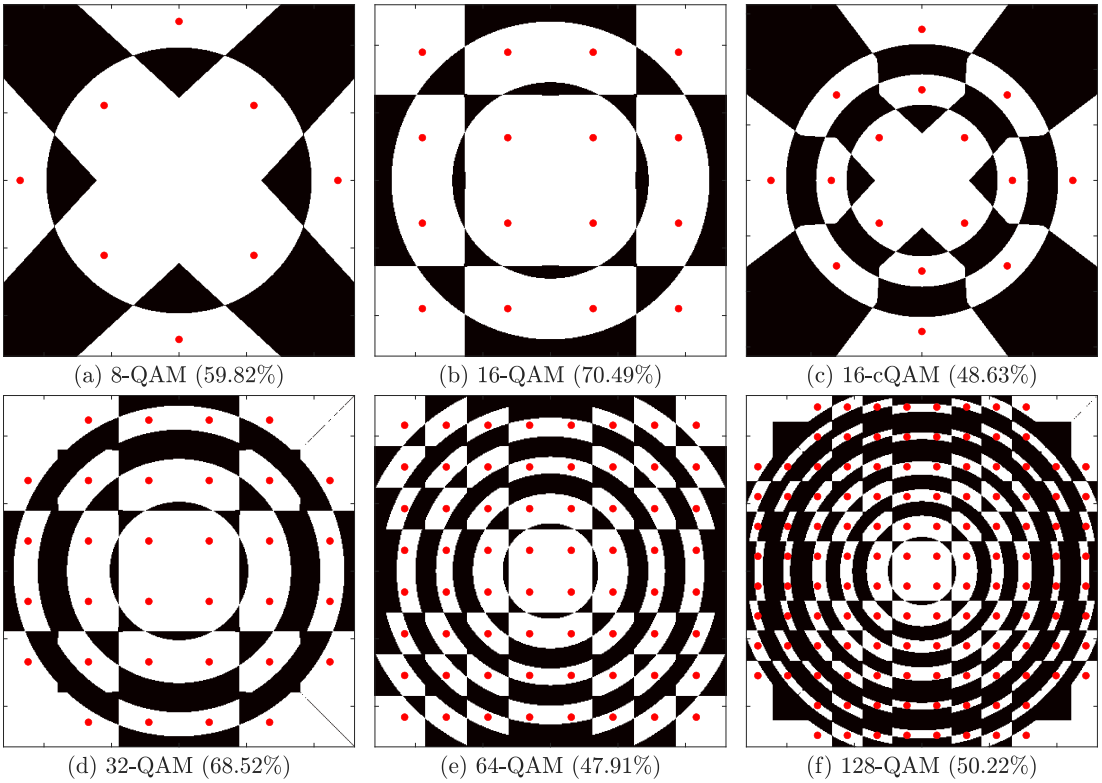
(d) $\eta = 5$ (e) $\eta \to \infty$ (f)

Figure 4.6 – Decision regions for 16-QAM and five values of η (a, b, c, d, and e). The common regions in (f) represent the sample spaces in which slicing can be performed by simple comparison (white area) or the squared distance equation needs to be actually computed (black area).

A further analysis of this common-region map is obtained and shown in Figure 4.7 for squared QAM constellations of order 8, 16, 32, 64, and 128, and a circular 16-QAM. For each case, it has been calculated the ratio of white area with respect to the total space and it is observed that in the worst case it represents only slightly below 50% and in the best case it is just above 70%.

4.2 Carrier phase estimation

Previously, in Chapter 2, the problem of carrier phase estimation (CPE) was briefly reviewed. The purpose of CPE (also known as carrier recovery, carrier tracking or phase-jitter suppression) is to follow the random fluctuations of the carrier phase evolution in order to recover the static modulated signal in which symbol



(a) 8-QAM (59.82%) (b) 16-QAM (70.49%) (c) 16-cQAM (48.63%)

(d) 32-QAM (68.52%) (e) 64-QAM (47.91%) (f) 128-QAM (50.22%)

Figure 4.7 – Computation of the ratio of common regions versus the total sample space (in percent), for QAM constellations of orders 8, 16, 32, 64, and 128 (a b, d, e, and f), and a circular 16-QAM (c).

detection can be performed (in this sentence, the term *static* refers to the fact that the recovered signal after CPE does not rotate in the complex plane and, therefore, the constellation appears to be still). To illustrate this, Figure 4.8 shows an example of how the signal constellation looks like at different stages of the DSP chain, namely, (a) the coherently received signal, (b) after equalisation, (c) after being compensated for any frequency offset, and (d) the recovered *still* constellation after CPE. The equalisation stage —from (a) to (b)— attempts to invert the effects of the channel transfer function in order to mitigate linear impairments such as CD, attenuation, reflections, etc. Regarding the stage from (b) to (c), in coherent systems it is necessary to either downconvert the signal from the *intermediate frequency* (IF, in heterodyne and intradyne systems), or correct for frequency fluctuations of the LO (homodyne systems), so the resulting signal is in baseband, where detection can be performed. In any case, the angular IF is expressed as the

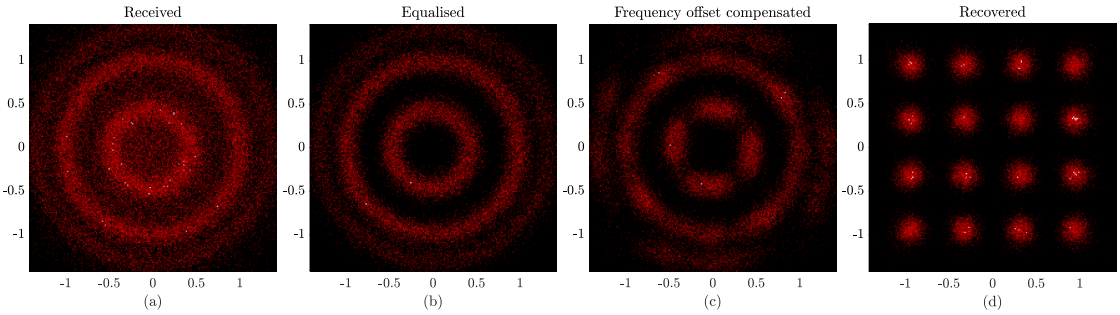


Figure 4.8 – Illustrative example of the aspect of the signal after different DSP stages in a coherent receiver: (a) received signal, (b) after equalisation, (c) after frequency offset compensation, and (d) after carrier phase estimation (CPE).

difference between the transmitter and LO optical centre angular frequencies, that is

$$\omega_{IF} = \omega_s - \omega_{LO} \tag{4.10}$$

and, if converted to Hz (i.e. $f_{IF} = \omega_{IF}/2\pi$), it can be of the order of few tens to hundreds of MHz and maybe even a few GHz depending on the sampling rate, the number of samples per symbol, and the baud rate [124]. Moreover, ω_{IF} can be first estimated ($\hat{\omega}_{IF}$) and the CPE block —from (c) to (d)— can deal with a small amount of residual frequency offset, apart from the phase noise itself [125].

Due to the exposed above, when only the CPE problem is considered, it is common to assume that the signal at the input of the CPE block has been already normalised, orthogonalised, equalised and compensated for the frequency offset that might have been present. Under this assumptions, the incoming signal samples received at each instant k are expressed as

$$y(k) = x(k) \exp[j\varphi(k)] + n(k) \tag{4.11}$$

where $x(k)$ are the transmitted symbols, $n(k)$ is complex AWGN which has zero mean and total variance N_0. The phase noise $\varphi(k)$ is a random process that results from the combined effect of both the transmitter and LO lasers linewidth. It is commonly modelled as a Wiener-Lévy process [126] (also known as a random walk) that can be described as follows

$$\varphi(k) = \varphi(k-1) + \psi(k) \tag{4.12}$$

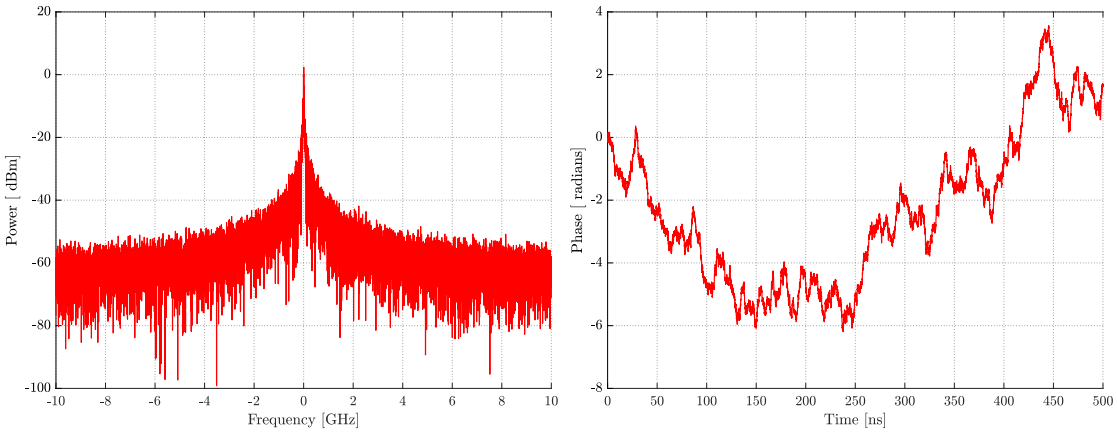


Figure 4.9 – Example of the Lorentzian shaped spectrum of a laser optical output (left) and the phase evolution governed by the Wiener-Lévy process or random walk (right)

where $\psi(k)$ is Gaussian distributed with zero mean and variance σ_p^2, which in turn can be expressed as

$$\sigma_p^2 = 2\pi\Delta\nu T_s \tag{4.13}$$

where $\Delta\nu$ is the sum of the full-width half-maximum (FWHM) linewidth of the transmitter and LO lasers (measured in Hz or seconds^{-1}), and T_s is the sampling period in seconds. From Equation (4.13) it can be seen that increasing the baud rate of a system will reduce the effective variance and, thus, improve its tolerance to the phase noise. Consequently, for benchmarking and performance evaluation purposes, it is common to express the amount of phase noise in terms of the dimensionless combined-linewidth and symbol-period product, $\Delta\nu T_s$. Figure 4.9 shows an example of the realisation of a phase noise process, with the Lorentzian spectrum on the left, and the *random walk* time evolution on the right.

In this section it is intended to present a way in which the approximated model presented in Chapter 3 can be of use in the CPE problem. For such purpose it is necessary to introduce two popular approaches, the maximum *a posteriori* (MAP) and the blind phase search (BPS) algorithms.

4.2.1 The maximum *a posteriori* algorithm

Given the channel described in Equation (4.11), the best estimate of the phase that can be obtained from the received samples $y(k)$ is the MAP estimate. There are two possible ways to perform the MAP estimation: the sequence and the symbol-by-symbol approaches [101]. In the first case, due to that the phase noise is a process with memory, the optimum phase estimate $\hat{\varphi}(k)$ is obtained as the sequence of values that maximize the probability $p(\mathbf{y}|\mathbf{x})$ where the boldfaced $\mathbf{y}$ and $\mathbf{x}$ indicate the vector containing samples that, ideally, extend the whole sequence of transmitted symbols. This is equivalent to minimising the log-probability function

$$\log(P) = \sum_k \left(\frac{|y(k) - \hat{x}(k) \exp[j\hat{\varphi}(k)]|^2}{2N_0} + \frac{(\hat{\varphi}(k) - \hat{\varphi}(k-1))^2}{2\sigma_p^2} \right) \tag{4.14}$$

where $\hat{x}(k)$ are symbol decisions performed during the execution and, therefore, it implies this is a decision-directed algorithm in which the phase noise and transmitted symbols are estimated jointly. It appears obvious that the sequence length requirement makes this an impracticable algorithm for implementation in any real system. Even using truncated sequences to a few-symbols long is already prohibitive in complexity for real-time systems [102].

Symbol-by-symbol approach

Alternatively, the MAP algorithm can be executed in a symbol-by-symbol fashion. When performed this way, it is assumed that previous estimations $\hat{\varphi}(k-1)$ (and $\hat{x}(k-1)$, too) are valid and therefore they no longer need to be considered in the sequence, which reduces to a single-sample long one, requiring the joint estimation of the most recent values of $\hat{\varphi}(k)$ and $\hat{d}(k)$ uniquely. If that is the case, the reduced log-probability function of Equation (4.14) becomes the one described by Equation (4.2), which indeed represents a simpler solution. Consequently, it makes sense to adopt the proposed metric D^2_{prop} for CPE purposes. The remainder of this section proposes how this can be done in a decision-directed CPE algorithm.

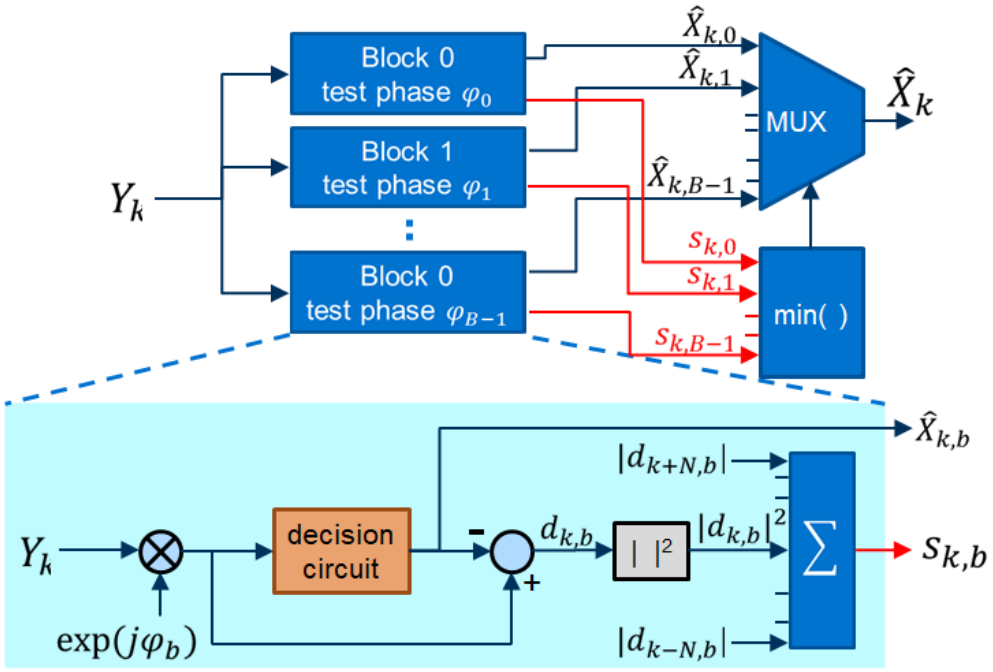


Figure 4.10 – Blind phase search algorithm hardware structure

4.2.2 The blind phase search algorithm

The blind phase search algorithm, or BPS, was first proposed by Pfau *et al.* in 2009 [127] and it was presented as a "hardware-efficient" concept for feedforward CPE with high-order QAM constellations. BPS has nowadays become the *de facto* reference algorithm among literature for CPE benchmarking due to its excellent performance and its relatively straightforward implementation. Although part of the BPS success relies on its high degree of parallelism in its hardware implementation, it is still somehow regarded as computationally expensive for real-time and on-chip implementation, due to the amount of resources consumed, as it will be seen.

Figure 4.10 depicts the block diagram of the BPS algorithm implemented in a feedforward fashion. Every received sample $y(k)$ is rotated by a number B of test angles θ_b where

$$\theta_b = \frac{b}{B} \cdot \Psi \tag{4.15}$$

for $b \in \{0, 1, \ldots, B-1\}$ and where Ψ is the constellation angle of symmetry and is generally equal to $\pi/2$ for conventional QAM formats (and this will be the value

considered throughout the text). The rotated samples $y_b(k) = y(k)\exp(j\theta_b)$ are sent to a decision block in which the slicing metric is based on minimum Euclidean distance, and which outputs the pair of values $\hat{x}_b(k)$ and $|d_b(k)|^2$, namely the decided symbol and the respective squared distance with respect to $y_b(k)$, that is

$$|d_b(k)|^2 = |y_b(k) - x_b(k)|^2 \tag{4.16}$$

In order to obtain one decoded symbol, the $2N+1$ distances corresponding to consecutive samples rotated by the same test angle θ_b are added up in a new variable $s_b(k)$ such as

$$s_b(k) = \sum_{n=-N}^{N} |d_b(k+n)|^2 \tag{4.17}$$

which is done in order to average out the noise contributions and where the optimum value of N depends on the amount of phase noise, being values between 6 and 10 a generally good choice. Similarly, the number of required test angles B depends on σ_p^2 as well as in the constellation order M, as denser constellations will demand a higher angular resolution.

Finally, the optimum joint estimate of phase and transmitted symbol is performed as follows: first, finding the minimum $s_b(k)$ implicitly solves for the phase estimate $\hat{\varphi}_b$, or what is the same, returning the index $b_{k,opt}$ as

$$b_{k,opt} = \arg\min_b s_b(k) \tag{4.18}$$

and, secondly, this index is used to control the multiplexer switch to output the estimated symbol $\hat{x}_b(k)$. It can be observed that because the phase and the transmitted symbol are estimated jointly, in a real system only the latter might be of interest, while the former can only serve for monitoring or performance evaluation.

4.2.3 A MAP-optimised BPS algorithm

The goal at this stage is to combine the best of both the MAP and BPS algorithms in order to develop a new solution that benefits from the statistical optimality of the first, and the more feasible hardware implementation of the latter. The proposed solution consists of replacing the Euclidean-based slicer and distance

computation and replace it by a block that performs pre-slicing by means of the common-region maps shown in Figure 4.7 and then compute the squared distance of the selected symbols using the proposed metric as

$$D^2_{\text{prop}}\left(y, \lfloor \hat{\mathbf{x}} \rfloor\right) = \Delta\rho^2 + \left(\gamma\Delta\varphi\right)^2 \tag{4.19}$$

$$= \left(|y| - |\lfloor \hat{\mathbf{x}} \rfloor|\right)^2 + \left(\eta + \frac{1}{|\lfloor \hat{\mathbf{x}} \rfloor|\,|y|}\right)\left(\arg\left(y\right) - \arg\left(\lfloor \hat{\mathbf{x}} \rfloor\right)\right)^2 \tag{4.20}$$

where $\lfloor \hat{\mathbf{x}} \rfloor$ is a vector containing either one pre-sliced symbol, in the best case, or a maximum of two symbols if y does not lay in a common region of the slicing map. When this is the case, the computation of Equation (4.20) will yield the minimum distance between the two, along with the decision index.

Moreover, also as a result of the research presented in this work, a different approach was followed in [128] to derive an equivalent metric to D^2_{prop}. The main difference posed by this alternative way is that apart from the squared distance calculation, it is possible, with an optional additional computation, to obtain an estimation of the angular component due to phase noise only, so it can be used in, for instance, phase noise variance estimation or phase noise signal reconstruction. This alternative approach is described as follows. Consider the example shown in Figure 4.11 (a) where the transmitted sample x has been posteriorly observed as the point y. The conventional BPS approach uses the Euclidean distance to compute D. Here, the aim is to consider D as the combination of linear and angular noise. Figure 4.11 (b) illustrates this, where x first suffers an angular deviation by an angle θ and a second linear displacement $\hat{D}$. Generally, the order in which these two transformations occur does not change the statistical behaviour of the system. If the samples x and y are given in polar coordinates, the observed angular difference

$$\Delta\varphi = \arg\left(y\right) - \arg\left(x\right) \qquad \Delta\varphi \in [-\pi, \pi) \tag{4.21}$$

$$= \alpha + \theta \tag{4.22}$$

where α is the angle shift incurred by the AWGN component $\hat{D}$, which can be written in polar coordinates according to the cosine rule of a triangle

$$\hat{D}^2 = |x|^2 + |y|^2 - 2\,|x|\,|y|\cos\left(\alpha\right) \tag{4.23}$$

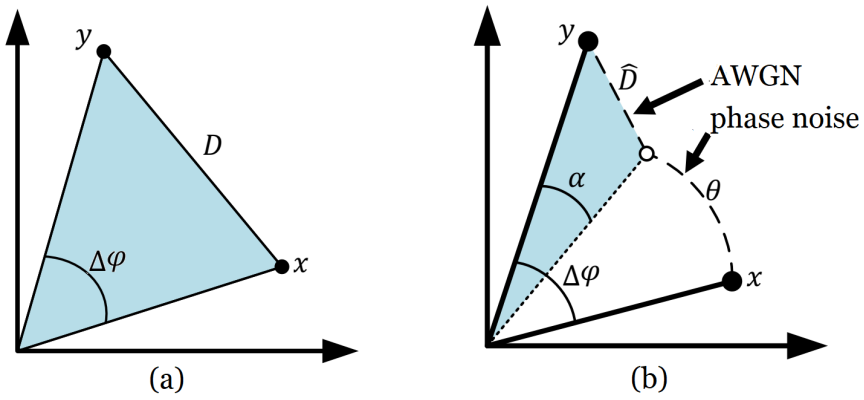


Figure 4.11 – (a) Example of using the phase-insensitive Euclidean distance versus (b) a model that splits the separation into AWGN and phase noise

and the log-P function given in Equation (4.2) can be rewritten in the form of the probabilistic distance

$$D^2 = D_{MAP}^2 = \frac{\hat{D}^2}{N_0} + \frac{\theta^2}{2\sigma_p^2} \tag{4.24}$$

$$= \frac{|x|^2 + |y|^2 - 2\,|x|\,|y|\cos(\alpha)}{N_0} + \frac{(\Delta\varphi - \alpha)^2}{2\sigma_p^2} \tag{4.25}$$

In order to find the minimum value of D_{MAP}^2 it is necessary to find the angle α_{opt} that satisfies the condition (since all the other variables are known)

$$\alpha_{opt} = \arg\min_{\alpha} D_{MAP}^2(\alpha)\,. \tag{4.26}$$

Differentiating with respect to α and setting it equal to zero yields

$$\frac{dD_{MAP}^2}{d\alpha} = \frac{2\,|x|\,|y|\sin(\alpha)}{N_0} - \frac{(\Delta\varphi - \alpha)}{\sigma_p^2} = 0 \tag{4.27}$$

where trying to solve for α leads to an expression of the form $\sin(\alpha) = a + b\alpha$ that would require a numerical method that iterates over α and that here is attempted to be avoided. To simplify this, it is possible to use the *small angle approximation* in Equation (4.27) by which $\sin(\alpha) \approx \alpha$ and then reduce to the simple expression

given by

$$\alpha_{opt} = \frac{\Delta\varphi}{\eta\mu + 1} \tag{4.28}$$

where it must be recalled that

$$\mu = |x|\,|y|$$

and η was defined in Equation (4.7). Next, the value of α obtained in Equation (4.28) can be used in Equation (4.25), where a further simplification can be done by replacing the cosine function by its Taylor expansion up to the second term, i.e. $\cos(\alpha) \approx 1 - \frac{\alpha^2}{2}$. The resulting expression is derived as follows

$$\begin{aligned}
D^2_{MAP} &\approx \frac{|x|^2 + |y|^2 - 2\,|x|\,|y|\left(1 - \frac{\alpha^2}{2}\right)}{N_0} + \frac{(\Delta\varphi - \alpha)^2}{2\sigma_p^2} \\
&= \frac{(|y| - |x|)^2 + |x|\,|y|\,\alpha^2}{N_0} + \frac{(\Delta\varphi - \alpha)^2}{2\sigma_p^2} \\
&= \frac{(|y| - |x|)^2 + \mu\left(\frac{\Delta\varphi}{\eta\mu+1}\right)^2}{N_0} + \frac{\left(\Delta\varphi - \left(\frac{\Delta\varphi}{\eta\mu+1}\right)\right)^2}{2\sigma_p^2} \qquad (\times N_0) \\
&= (|y| - |x|)^2 + \mu\left(\frac{\Delta\varphi}{\eta\mu + 1}\right)^2 + \frac{\Delta\varphi^2\left(\frac{\eta\mu}{\eta\mu+1}\right)^2}{\eta} \\
&= (|y| - |x|)^2 + \left(\frac{\mu}{(\eta\mu+1)^2} + \frac{\eta\mu^2}{(\eta\mu+1)^2}\right)\Delta\varphi^2 \\
&= (|y| - |x|)^2 + \left(\eta + \mu^{-1}\right)^{-1}\Delta\varphi^2
\end{aligned} \tag{4.29}$$

which is equivalent to D^2_{prop} in Equation (4.6) and that in [128] was coined as D^2_{pMAP} for "polar MAP" distance. It has been thus shown that following either of the approaches leads to the same result. Figure 4.12 illustrates the modified BPS hardware structure where the Euclidean slicer and distance calculation has been replaced by the Cartesian-polar hybrid slicer and the polar metric D^2_{prop}.

4.2.4 Simulation Results

A set of simulations have been carried out in order to evaluate the performance of the proposed MAP-optimised BPS algorithm in the problem of CPE. The tested modulation formats are square QAM of orders 4, 16, 64, and 256. In order to avoid

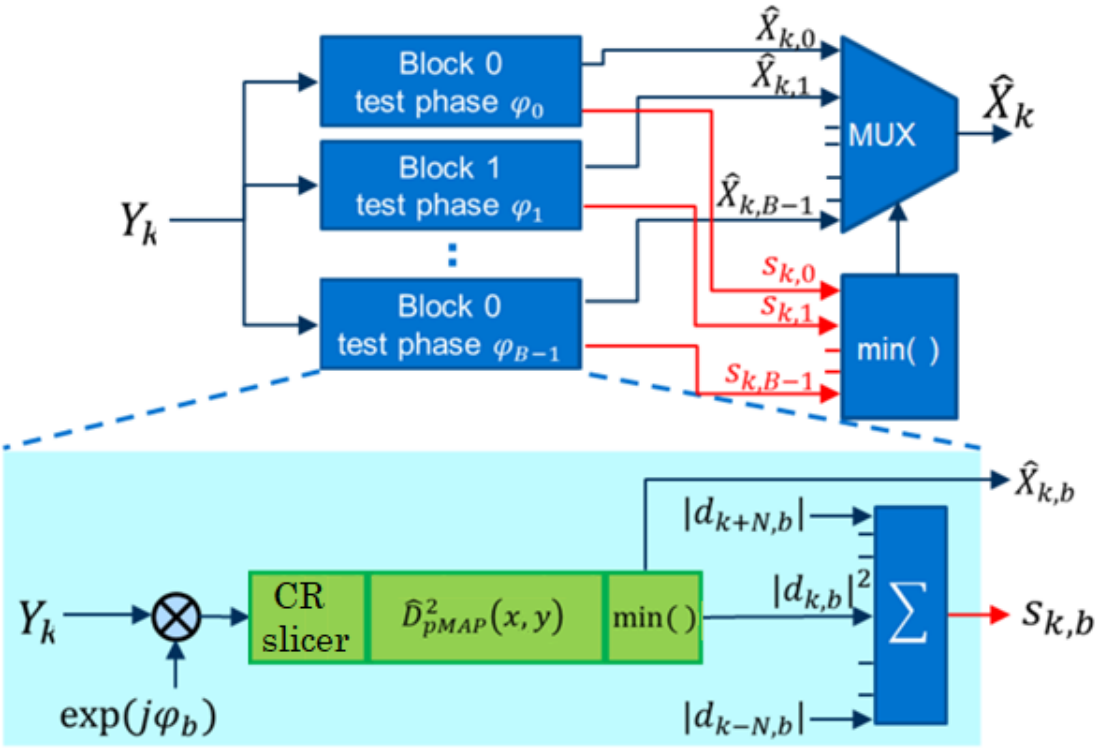


Figure 4.12 – The proposed modified BPS structure replaces the Euclidean detector by the proposed common-region (CR) slicer followed by the proposed metric $\hat{D}^2_{pMAP}$

burst errors due to cycle slips, and to avoid the 4-fold symmetry ambiguity of the estimated phase in square QAM constellations, differential coding has been applied for the bit-to-symbol mapping [129]. It suffices to encode the two bits that identify the quadrant using Gray coding. The decoding process, as presented in [130], is described as follows

$$n_{o,k} = (n_{r,k} - n_{r,k-1} + n_{j,k}) \mod 4 \qquad n_{o,k}, n_{r,k}, n_{j,k} \in \{0,1,2,3\} \tag{4.30}$$

where $n_{o,k}, n_{r,k}, n_{j,k}$ are the differentially decoded quadrant, received quadrant, and jump numbers, respectively. A slight modification with respect to [130] is required in order to assign the detected quadrant jumps, given as

$$n_{j,k} = \begin{cases} 1, & \text{if } b_{k,opt} - b_{k-1,opt} > B/2 \\ 3, & \text{if } b_{k,opt} - b_{k-1,opt} < -B/2 \\ 0, & \text{otherwise} \end{cases} \tag{4.31}$$

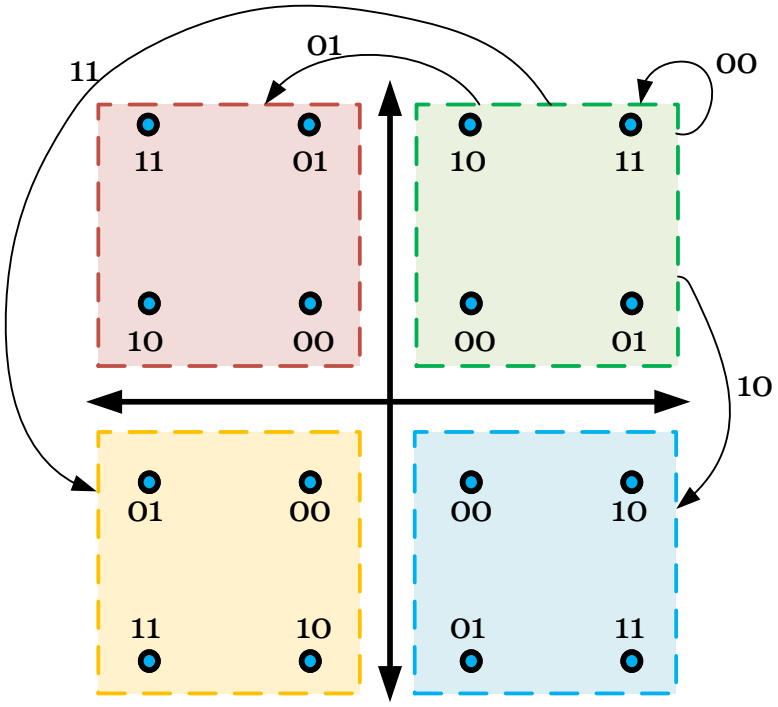


Figure 4.13 – Example of differential bit coding for 16-QAM as used in all the simulations. The differential encoding is similarly applied to higher-order QAM formats using the two most significant bits (MSB) for each quadrant

where $b_{k,opt}$ is the index returned by the minimum distance selector and given in Equation (4.18). An example of a differentially encoded 16-QAM is illustrated in Figure 4.13.

Consequently, the symbol mapping resulting from the encoding process is no longer Gray-coded, yielding a penalty that depends on the modulation format. Particularly for QPSK, the result of differential decoding is that a symbol error will always be followed by another symbol error, therefore duplicating the bit-error count. As the QAM order M increases, the penalty decreases, being practically 0 dB for high-order constellations [129, 131].

In the simulation, a large amount of data has been generated according to the channel described by Equation (4.11), using a wide range of SNR and phase noise values. For each tested scenario, a sequence of $2 \cdot 10^5$ symbols is randomly

generated, where the filter length is set to $L = 2N + 1$, $N = 5$, and the number of test angles is $B = 32$. What it has been measured is the SNR penalty with respect to the theoretically achievable sensitivity in the absence of phase noise, also with the differential coding explained earlier for a fair comparison. This achievable sensitivity is targeted to two values of HD pre-FEC BER values, namely 10^{-3} and 10^{-2}.

Figure 4.14 shows the results for the first targeted BER of 10^{-3}. The SNR penalty in dB is given for the conventional BPS using the Euclidean distance (red line, pointing-down triangle marker) and for the modified BPS with the polar metric (blue line, pointing-up triangle marker). It can be seen that whilst there is no difference when using QPSK (due to the 4-fold symmetry of the constellation and having only one symbol per quadrant), there is a clear improvement in the linewidth tolerance which increases with the constellation order. Effectively, the polar metric enables to push the maximum tolerable value of $\Delta\nu T_s$. Additionally, Table 4.1 summarises the maximum tolerable values of $\Delta\nu T_s$ at the 1-dB SNR penalty point for the BER of 10^{-3}, and the maximum tolerable combined linewidth $\Delta\nu$ if it was to be used in a system operating at 32 GBaud.

Similarly, Figure 4.15 shows the same performance results, but this time being targeted to a BER of 10^{-2}, and Table 4.2 shows the same summary of maximum values of phase noise as its counterpart, but for the new value of BER = 10^{-2}.

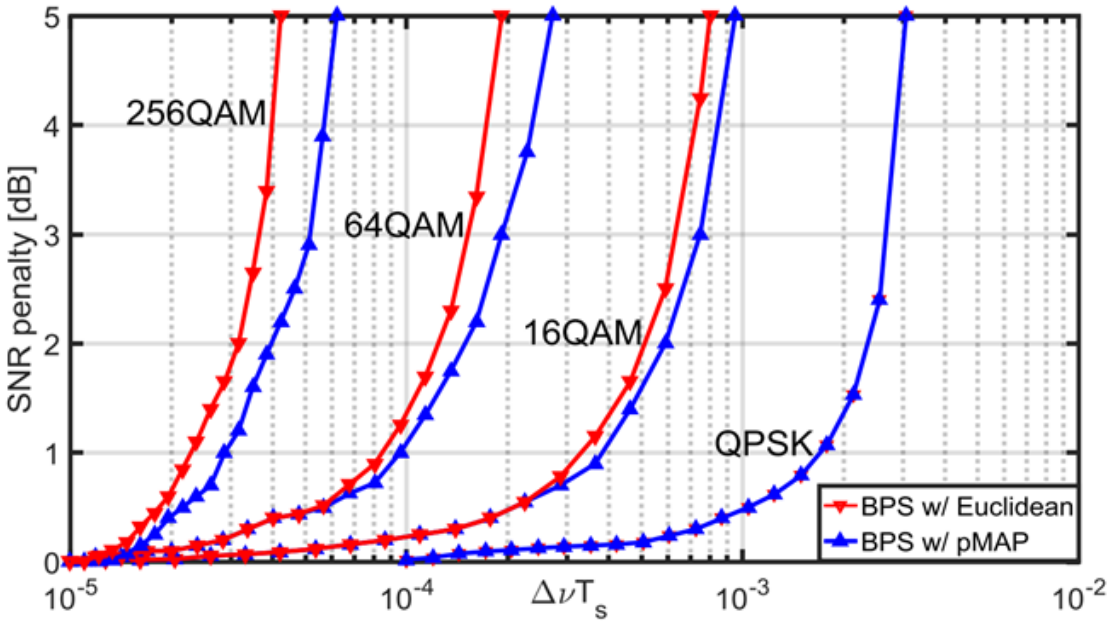


Figure 4.14 – SNR penalty in dB with respect to the theoretically achievable sensitivity for a target pre-FEC BER of 10^{-3} and for differentially encoded square QAM constellations of order 4, 16, 64, and 256. Results are obtained with the conventional (Euclidean) BPS (red line, ∇ marker) and the MAP-optimised BPS (D^2_{pMAP}, blue line, Δ marker)

Table 4.1 – Phase noise tolerance comparison between Euclidean versus optimised BPS and maximum tolerable linewidths in a 32 GBaud system at a BER of 10^{-3}

Constellation	max $\Delta\nu T_s$		max $\Delta\nu$ @ 32 GBaud	
	Euclidean	D^2_{prop}	Euclidean	D^2_{prop}
QPSK	$1.7 \cdot 10^{-3}$		54 MHz	
16-QAM	$3.3 \cdot 10^{-4}$	$3.8 \cdot 10^{-4}$	10.6 MHz	12.2 MHz (+15%)
64-QAM	$0.8 \cdot 10^{-4}$	$1 \cdot 10^{-4}$	2.6 MHz	3.2 MHz (+25%)
256-QAM	$2.3 \cdot 10^{-5}$	$3 \cdot 10^{-5}$	736 kHz	960 kHz (+30%)

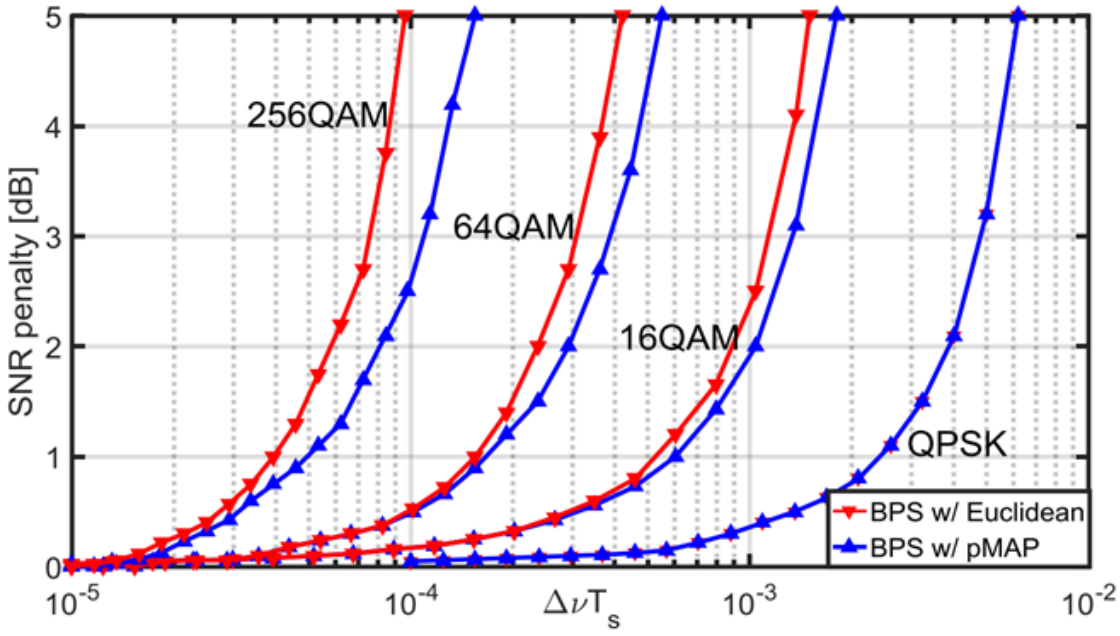


Figure 4.15 – SNR penalty in dB with respect to the theoretically achievable sensitivity for a target pre-FEC BER of 10^{-2} and for differentially encoded square QAM constellations of order 4, 16, 64, and 256. Results are obtained with the conventional (Euclidean) BPS (red line, ∇ marker) and the MAP-optimised BPS (D^2_{pMAP}, blue line, Δ marker)

Table 4.2 – Phase noise tolerance comparison between Euclidean versus optimised BPS and maximum tolerable linewidths in a 32 GBaud system at a BER of 10^{-2}

Constellation	max $\Delta\nu T_s$		max $\Delta\nu$ @ 32 GBaud	
	Euclidean	D^2_{prop}	Euclidean	D^2_{prop}
QPSK	$2.5 \cdot 10^{-3}$		80 MHz	
16-QAM	$5 \cdot 10^{-4}$	$6 \cdot 10^{-4}$	16 MHz	19.2 MHz (+20%)
64-QAM	$1.5 \cdot 10^{-4}$	$1.85 \cdot 10^{-4}$	4.8 MHz	5.9 MHz (+23%)
256-QAM	$3.9 \cdot 10^{-5}$	$5 \cdot 10^{-5}$	1.25 MHz	1.6 MHz (+28%)

4.3 Symbol error rate estimation

Another possible application of the PCAWGN channel model is that of HD symbol error rate (HD-SER) or, analogously, bit-to-error ratio (HD-BER) estimation. Many works found in the literature have dealt with the problem of SER (e.g. in [132]) and BER (e.g. in [133]) estimation, however, they often assume that the slicing function at the receiver is Euclidean by default. In this section it is presented an alternative approach that considers, instead, the metric given by D^2_{prop}.

Due to the fact that in the presence of angular noise the optimum symbol decision boundaries are not straight lines it is necessary to divide the sample space into a grid of discrete sections over which to integrate the symbol probability. Since the proposed PCAWGN model is defined in polar coordinates, it will be convenient to build such a grid likewise.

4.3.1 Polar grid and transmitted probability matrix

The procedure to build the polar grid is as follows: build two vectors of equally spaced moduli and angles as (see Figure 4.16(a))

$$\mathbf{r} = [0, r_2, \ldots, r_{N_r}, R_{lim}] \tag{4.32}$$

$$\mathbf{\Phi} = \left[-\pi, \phi_2, \ldots, \phi_{N_\phi}, \pi\right] \tag{4.33}$$

where R_{lim} must be selected so the tail probability $p\left(R_{lim} > R_{\max}\right) > \epsilon$, and where $R_{\max}$ is the largest modulus in the constellation. Compute each *transmitted probability* matrix $\mathbf{T}_m$ of size $(N_r,\, N_\phi)$ for all the $m = 1, \ldots, M$ symbols in the constellation $\mathcal{X}$ as:

$$\mathbf{T}_m\left(u, v\right) = p\left(\Delta r_{u,m}, \frac{N_0}{2}\right) \cdot p\left(\Delta\phi_{v,m}, \sigma^2_{\phi,u,m}\right) \tag{4.34}$$

where

$$\Delta r_{u,m} = \mathbf{r}\left(u\right) - \left|\mathcal{X}\left(m\right)\right|, \qquad u = 1, \ldots, N_r \tag{4.35}$$

$$\Delta\phi_{v,m} = \mathbf{\Phi}\left(v\right) - \arg\left(\mathcal{X}\left(m\right)\right), \qquad v = 1, \ldots, N_\phi \tag{4.36}$$

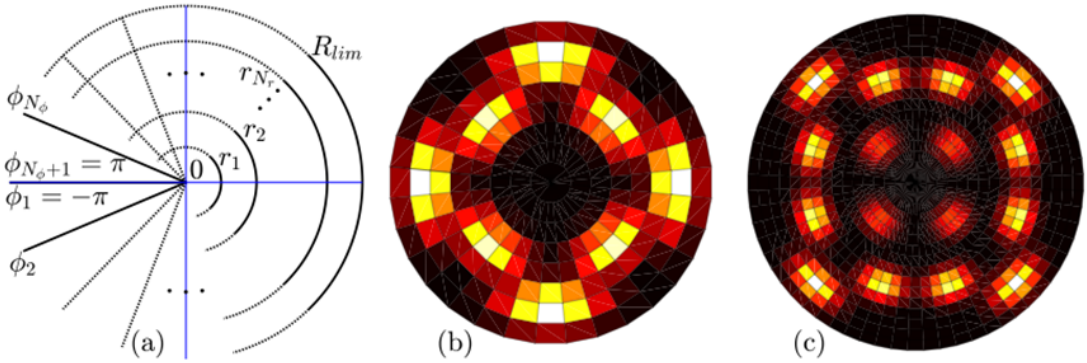


Figure 4.16 – (a) Example of the polar grid construction with equispaced angles and radii, (b) and (c) lower- and higher-resolution probability density maps for arbitrary amounts of AWGN and phase noise, in 8- and 16-QAM constellations, respectively.

and where $\sigma_{\phi,u,m}$ is the radius-dependent variance

$$\sigma^2_{\phi,u,m} = \frac{N_0}{2r'_u \left|\mathcal{X}\left(m\right)\right|} + \sigma^2_p, \quad r'_u = \frac{r_u + r_{u+1}}{2} \tag{4.37}$$

The integral of the probability $p\left(\cdot\right)$ was computed by finding the differences of its cumulative density function (F), that is

$$p\left(\Delta r_{u,m}, \frac{N_0}{2}\right) = \mathrm{F}\left(\Delta r_{u+1,m}, \frac{N_0}{2}\right) - \mathrm{F}\left(\Delta r_{u,m}, \frac{N_0}{2}\right) \tag{4.38}$$

$$p\left(\Delta \phi_{v,m}, \sigma^2_{\phi,u,m}\right) = \mathrm{F}\left(\Delta \phi_{v+1,m}, \sigma^2_{\phi,u,m}\right) - \mathrm{F}\left(\Delta \phi_{v,m}, \sigma^2_{\phi,u,m}\right) \tag{4.39}$$

where F() is defined based on the *error function*[1] as

$$F\left(a, \sigma\right) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{a}{\sigma\sqrt{2}}\right)\right] \tag{4.40}$$

Figure 4.16 (b) and (c) show two examples of the computed $\mathbf{T}$ for 8- and 16-QAM.

[1]for simulations it was used the MATLAB r2017b implementation of `erf()`

4.3.2 Computation of symbol error probability

Similarly to the transmitted probabiity matrix, build a decision matrix $\mathbf{D}$ containing the natural indexes (1, ..., M) of the symbols that minimize the slicing function metric (any chosen one) with respect to the points in the polar grid, as

$$\mathbf{D}(u,v) = \arg \min_{m \in \{1,\ldots,M\}} \left| \left(\mathbf{r}(u), \mathbf{\Phi}(v)\right) - \mathcal{X}(m) \right|^2 \tag{4.41}$$

A symbol-to-symbol probability map of size $M \times M$, containing the probability that the transmitted symbol $\mathcal{X}(x)$ will be received as the symbol $\mathcal{X}(y)$, can be computed as

$$\mathbf{S}_{x,y} = \sum_{u,v} \mathbf{T}_x(u,v) \cdot \mathbb{1}_y\left(\mathbf{D}(u,v)\right) \tag{4.42}$$

with $x, y \in 1, \ldots, M$, and where the identification function $\mathbb{1}_y$ is equal to 1 when $w = y$ and 0 otherwise. The total BER is obtained as

$$BER = p(x_m) \sum_{x,y=1}^{M} \mathbf{S}_{x,y} \mathrm{e}_{x,y} \tag{4.43}$$

where $\mathrm{e}_{x,y}$ is the number of bit differences (errors) between symbols $\mathcal{X}(x)$ and $\mathcal{X}(y)$. Notice that in Equation (4.43) $p(x_m)$ can be replaced by $\frac{1}{M}$ if all the symbols x_m are transmitted with equal probability.

4.3.3 Simulation results

An important parameter that needs to be chosen and that has a significant impact in the performance of the estimation is the size of the polar grid, i.e. the number of moduli and angles that form the vectors $\mathbf{r}$ and $\mathbf{\Phi}$ (N_r and N_ϕ, respectively). In this aspect, two different cases have been considered and that are shown in Table 4.3: first, a maximum resolution scenario where the size of the grid is 256 for both N_r and N_ϕ, that is, with 8 bits of resolution each. From there, a lower resolution scenario has been investigated for each constellation order, where the resolution was decreased until the estimated BER was, in the worst case, within the 95% with respect to the maximum resolution scenario.

Table 4.3 – Sets of modulus and angle grid resolutions

	High resolution (# bits)		Low resolution (# bits)	
QAM	N_r	N_ϕ	N_r	N_ϕ
8	256 (8)	256 (8)	8 (3)	32 (5)
16			16 (4)	64 (6)
32			32 (5)	64 (6)
64			32 (5)	64 (6)
128			32 (5)	64 (6)
256			64 (6)	64 (6)

Figure 4.17 shows the predicted BER using the high-resolution (dotted red line) and the low-resolution grid (dashed blue line) against Monte Carlo simulations of 2^{20} symbols per data point (circle markers), for 8-, 16-, 32-, 64-, 128-, and 256-QAM constellations loaded with AWGN (given by SNR in dB) and Gaussian phase noise (given by σ_p^2). For each case, the maximum modulus R_{lim} is found using $\epsilon = 0.001$ and the decisions in $\mathbf{D}$ are made using the polar metric D^2_{prop}, although using the Euclidean distance did not report any degradation on estimation performance. As it can be seen, the estimated BER curves exhibit excellent agreement with the numerical simulations for all the applied combinations of noise in all scenarios, being practically indistinguishable the difference between high and low resolution grids.

4.4 Blind estimation of SNR and phase noise

In the channel model approximation derived in Chapter 3, as well as in the applications described in the present chapter, it was always assumed that the noise parameters were known. However, in many scenarios this information might *a priori* not be known, impeding or limiting the use of the mentioned metrics and applications. Because the amount of noise accumulated during transmission and present in the received signal is not something that the transmitter would know about, it is desirable to have techniques that can be performed at the receiver

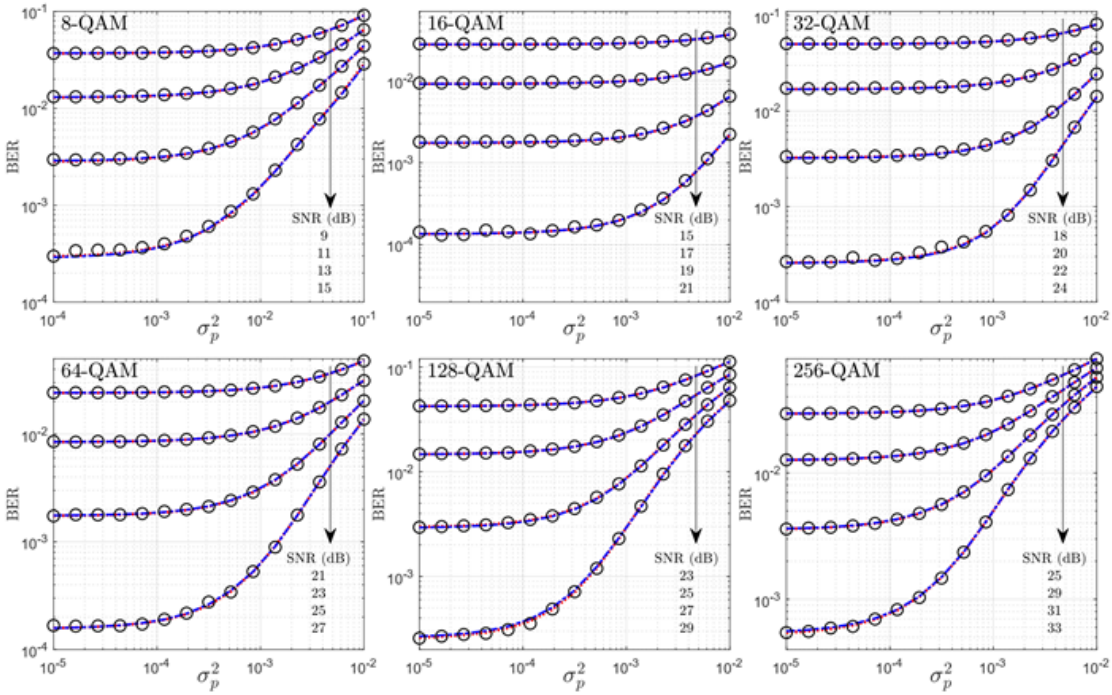


Figure 4.17 – BER estimation curves obtained with the proposed method using high-resolution grid (dotted red line) and low-resolution grid sets given in Table Table 4.3(dashed blue line). Results are shown for QAM constellations of order $M = 8,\ 16,\ 32,\ 64,\ 128,$ and 256, loaded with phase noise (15 values of σ_p^2 and AWGN (4 values of SNR in each scenario). For comparison, curves are shown along with results obtained by Monte Carlo simulation (circles).

in order to estimate how much noise is contained in the signal. This problem is commonly known as SNR estimation, and methods like those based in monitoring the *error vector magnitude* (EVM) are popular in decision-directed schemes [134], while those based in statistical properties and high-order moments (HOM) [135] are the non data-aided counterparts. However, these methods are mainly targeted to AWGN only and the problem posed in this work requires the estimation of phase noise too. In what follows, a blind joint estimation technique is proposed and evaluated.

Consider again the channel model given by Equation (4.11) seen before in Section 4.2 for the study of the CPE problem, and reproduced here

$$y(k) = x_m(k) \exp[j\varphi(k)] + n(k)$$

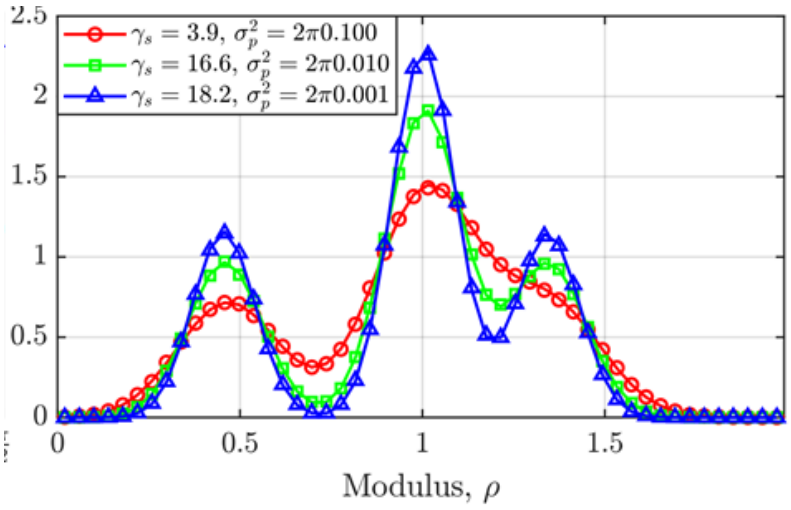


Figure 4.18 – Comparison of moduli distributions from simulated data (markers) and the PDF model (solid lines). Curves are shown for a 16-QAM constellation, histograms obtained with 10^7 symbols and the SNR is given by γ_s in dB

where the symbols $x_m(k)$ are drawn from the constellation $\mathcal{X} = \{x_1, \ldots, x_M\}$ of size M. In order to get rid of the channel memory imposed by the phase noise process, it is convenient to assume the receiver operates with differential detection, by which the observed variables are the received modulus ρ and the angle difference with respect to the last received symbols, $\Delta\theta = \arg[y(k)] - \arg[y(k-1)]$. Therefore, for a given constellation and assuming known noise variances, N_0 and σ_p^2, the joint PDF $p(\rho, \Delta\theta)$ is approximated as

$$p(\rho, \Delta\theta) \approx p_\rho(\rho) \cdot p_\theta(\Delta\theta) \tag{4.44}$$

where

$$p_\rho(\rho) = \sum_{m=1}^{M} p(x_m) \frac{\rho}{N_0/2} \exp\left(-\frac{\rho^2 + |x_m|^2}{N_0}\right) I_0\left(\frac{\rho|x_m|}{N_0/2}\right) \tag{4.45}$$

$$p_\theta(\Delta\theta) = \sum_{u,v=1}^{M} p(\xi_{u,v}) \frac{1}{\sqrt{2\pi\sigma_\theta^2}} \exp\left(-\frac{\Delta\theta^2}{2\sigma_\theta^2}\right) \tag{4.46}$$

and where $\xi_{u,v}$ represents a symbol transition $\xi_{u,v} := (x_u \to x_v)$, for any $u, v \in \{1, \ldots, M\}$ and where the angular variance is decomposed in the AWGN and phase

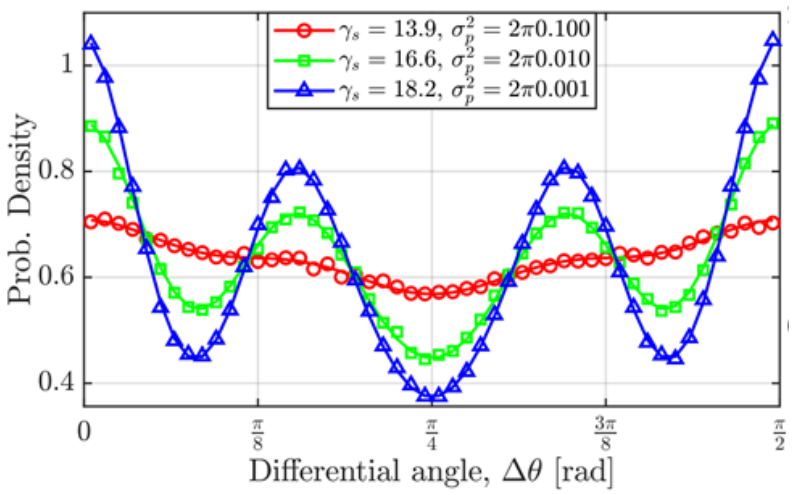


Figure 4.19 – Comparison of differential angle distributions from simulated data (markers) and the PDF model (solid lines). Curves are shown for a 16-QAM constellation, histograms obtained with 10^7 symbols and the SNR is given by γ_s in dB

noise contributions as

$$\sigma_\theta^2 = \frac{N_0/2}{|x_u|\,|x_v|} + \sigma_p^2 \tag{4.47}$$

Assuming that symbols are equiprobable, then $p(x_m) = M^{-1}$ and $p(\xi_{u,v}) = M^{-2}$.

Given a set of L received samples contained in the vector $\mathbf{y}$, compute the normalised histograms of their moduli, $\mathbf{h}_r = \overline{\text{hist}}(|\mathbf{y}|, \mathbf{r})$ and their modulo-$\frac{\pi}{2}$ differential angles, $\mathbf{h}_a = \overline{\text{hist}}(\arg(\mathbf{y}(\mathbf{k}+\mathbf{1})), \arg(y(k)), \mathbf{a})$ where $k = 1, \ldots, L-1$ and $\mathbf{r}$ and $\mathbf{a}$ are the bin vectors with N_r and N_a bins, respectively. Figure 4.18 shows the normalised histogram $\mathbf{h}_r$ of moduli from simulated data (markers) versus the probability density function obtained by Equation (4.45) (solid lines). Similarly, Figure 4.19 shows the normalised histogram $\mathbf{h}_a$ of differential angles from simulated data (markers) versus the probability density function obtained by Equation (4.46) (solid lines); both cases are for a 16-QAM constellation and for three values of SNR (given by γ_S in dB) and phase noise. The joint estimation of AWGN and phase noise variances can be performed by finding the pair $(\hat{N}_0, \hat{\sigma}_p^2)$ that minimises the sum of squared errors between the histograms and the model, that is

$$\left(\hat{N}_0, \hat{\sigma}_p^2\right) = \arg\min_{N_0, \sigma_p^2} \left[\sum^{N_r} (\mathbf{h}_r - p_\rho(\mathbf{r}))^2 + \sum^{N_a} (\mathbf{h}_a - p_\theta(\mathbf{a}))^2\right] \tag{4.48}$$

Table 4.4 – Test Cases for AWGN (γ_s, in dB) and PN (σ_p^2) estimation

Case	Parameter	16-QAM	32-QAM	64-QAM	128-QAM	256-QAM
I	γ_s	13.9	17.8	20.5	23.7	29.6
	$\langle\widehat{\gamma_s}\rangle$	13.89	17.79	20.48	23.67	29.48
	std($\widehat{\gamma_s}$)	1.2e-3	1.7e-3	1.9e-3	2e-3	4.1e-3
II	γ_s	16.6	20.1	23	26.1	32
	$\langle\widehat{\gamma_s}\rangle$	16.59	20.09	22.97	26.06	31.8
	std($\widehat{\gamma_s}$)	0.9e-3	1.1e-3	1.3e-3	2.3e-3	1.8e-3
I & II	σ_p^2	$-5 \leq \log_{10}\left(2\pi\sigma_p^2\right) \leq -2$				
III & IV	γ_s	13 . . . 20	16 . . . 23	19 . . . 26	22 . . . 29	27 . . . 34

4.4.1 Simulation results

The proposed method for blind joint estimation of AWGN and phase noise has been evaluated on QAM constellations of orders M=16, 32, 64, 128, and 256. In all cases, the normalised histograms $\mathbf{h}_r$ and $\mathbf{h}_a$ have been computed with $N_r = N_a = 128$ bins. Each resulting data point has been simulated 100 times for the acquisition of reliable values of mean and standard deviation. For each repetition, a total of $L = 10^6$ symbols have been generated. Two first cases are defined, Case I and Case II, where σ_p^2 is swept over a range of 30 values while the SNR per symbol (γ_S) is fixed to either one of two values, namely targeting HD-BER values of 10^{-2} and 10^{-3}, respectively. The values applied are shown in Table 4.4 along with the mean $\langle\hat{\gamma}_S\rangle$ and standard deviation std($\hat{\gamma}_S$) of the estimated SNR. In the last two Cases, III and IV, γ_S is swept to take 8 values within different ranges for each constellation order, while σ_p^2 is either zero (Case III) or equal to the applied N_0 in each scenario (Case IV). The estimated $\hat{\sigma}_p^2$ for Case I and II is shown in the error-bar plots shown in Figure 4.20, while the estimated SNR ($\hat{\gamma}_S$) for all scenarios in Cases III and IV is shown in Figure 4.21. In all cases, an excellent agreement between applied and estimated values is observed.

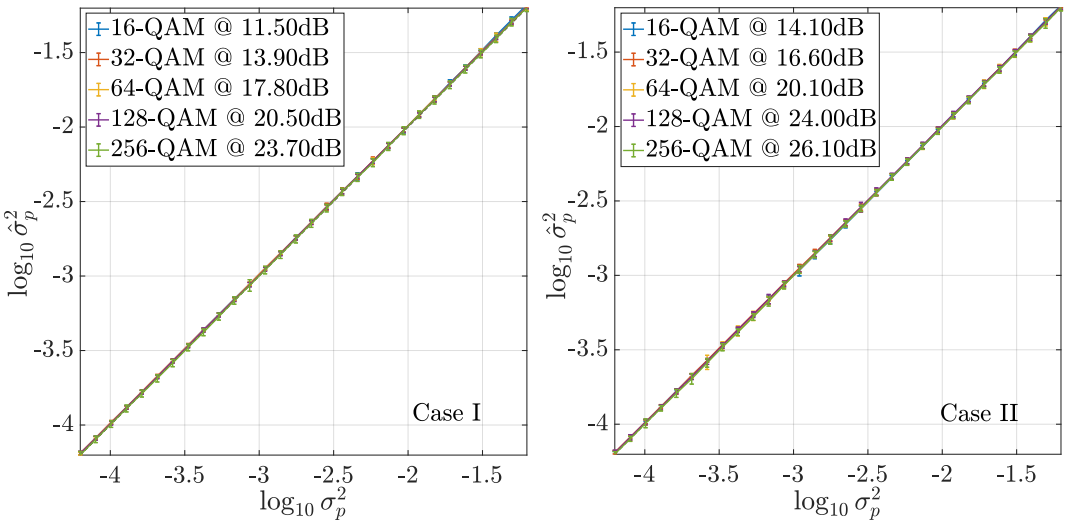


Figure 4.20 – Error-bar plot showing the phase noise estimation results for Cases I (left, BER $= 10^{-2}$) and II (right, BER $= 10^{-3}$), for five QAM constellations of order 16, 32, 64, 128, and 256, and where for each case the SNR has been kept constant.

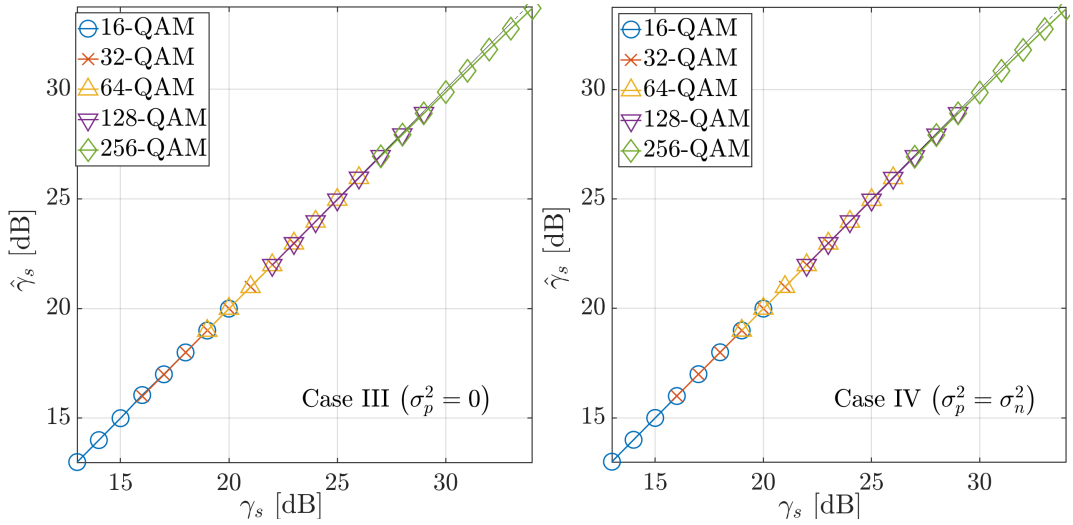


Figure 4.21 – Error-bar plot showing the SNR estimation results for Cases III (left) and IV (right), for five QAM constellations of order 16, 32, 64, 128, and 256.

To give a deeper insight, Figure 4.22 illustrates the worst-case relative error (%) for the estimation of σ_p^2 in Cases I and II while Figure 4.23 exposes the relative error for the estimated γ_S in Cases III and IV. In the first case, for σ_p^2, the relative error is always below 4% (although for most of the tested values and constellations it is below 3%), except for the higher values of σ_p^2 with 16-QAM, where it becomes

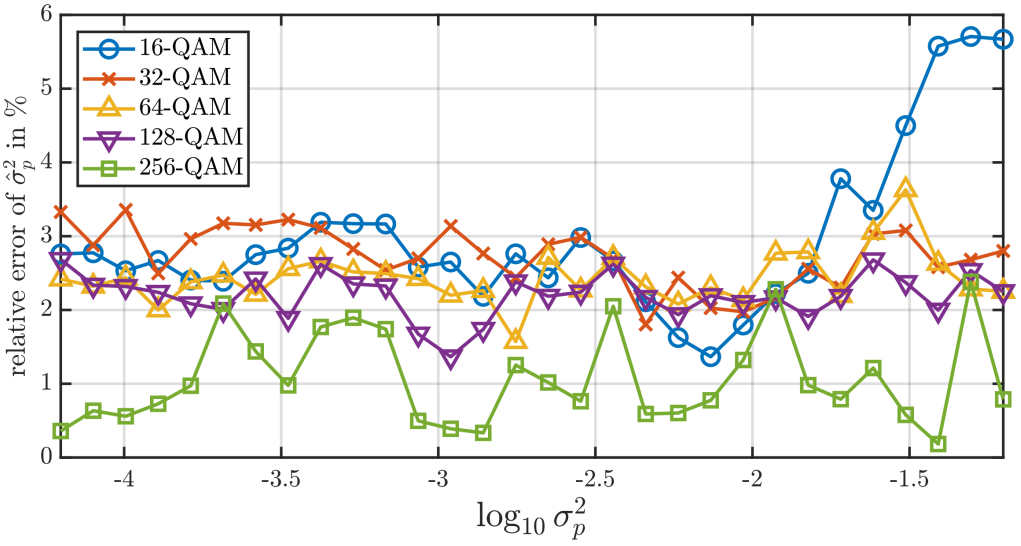


Figure 4.22 – Worst-case relative error for the estimated $\hat{\sigma}_p^2$ in Cases I and II.

just above 5%, while in the second cases, for $\hat{\gamma}_S$, the relative error stays below 1%.

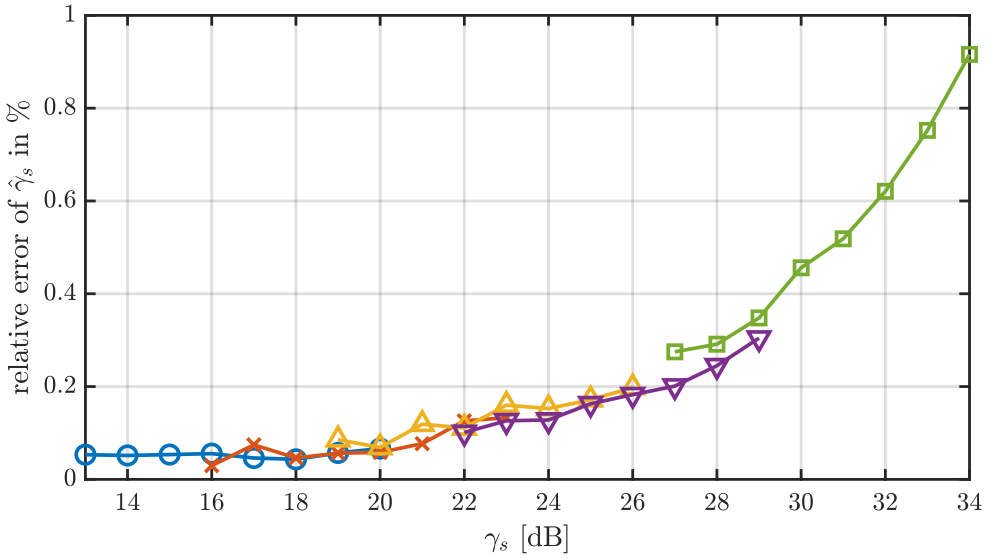


Figure 4.23 – Worst-case relative error for the estimated $\hat{\gamma}_S$ in Cases III and IV.

4.5 Conclusions

This chapter has been dedicated to explore and evaluate some potential applications where the proposed approximated model of the PCAWGN can be reliably used to improve the phase-noise tolerance of existing techniques.

In particular, from the receiver side, it has been tested the model as slicing function, clearly outperforming the Euclidean slicer in the presence of phase noise. Similarly, the reference algorithm for CPE, the BPS algorithm, has been shown to improve its linewidth tolerance if it is appropriately modified to include the proposed approximated model. Finally, a method for the blind joint estimation of AWGN and phase noise has been presented and tested, with quite successful results.

Although not fully covered in this chapter, but suggested in Chapter 6, such an extracted information about the noise present in the channel, could be used at the transmitter side in order to adapt aspects like, for instance, the constellation shape (geometric or probabilistic) in order to maximise the channel capacity in the detected scenario.

Chapter 5

On the complex modulation of DFB lasers

5.1 Introduction

While in the past the frequency chirp produced by lasers under direct intensity modulation has been regarded as an undesired effect —due to the pulse broadening caused by chromatic dispersion—, it has more recently attracted the attention of researchers who have shown that when using a coherent receiver, having access to the phase information enables a significant improvement against its direct-detection counterpart. This technique has been coined as *complex modulation* (CM) of directly modulated lasers (DML), or simply CM-DML.

However, the part "complex" of the term could be misleading if one understands that it implies the ability to modulate, *independently*, the real and imaginary parts of the electric field. In reality, DMLs are driven by a varying injected current, I, that produces a proportionally[1] varying optical power $P \propto I$ which in turn leads to a frequency modulation or *chirp*, Δf (blue-shift for a power increment, i.e. positive ΔP, red-shift for power decrement, i.e. negative ΔP). Similarly, the part "modulation" of the term could be arguable more accurate to be referred to as "demodulation", since the previous works on CM-DML only take into account the phase information at the receiver side, whilst at the transmitter the laser is modulated in a conventional way with the well-known PAM4 constellation. Nevertheless,

[1] Assuming the laser is DC-biased to operate in the linear region

there is no doubt about the substantial improvement offered so far by the CM-DML technique and the aim of this chapter is to explore the possibility of extending its capabilities by using the concepts developed in the previous chapters of this thesis.

While the previous chapter already introduced some of the potential applications of the approximated PCAWGN channel model described in Chapter 3, the present chapter introduces its potential application in CM-DML. The justification of its separation in a new chapter is that, although the initial research was carried out to investigate the benefits of using the proposed model with the current state-of-the-art CM-DML techniques, a further extension to the work has been implemented, bringing the concept to the transmitter side where the simultaneous amplitude and phase modulation of DML is exploited in order to design new constellations that could improve the channel capacity in cost-sensitive coherent systems.

This chapter is structured as follows: a theoretical introduction to the chirp effects in DFB lasers, necessary to understand the principles of CM-DML, is given in Section 5.2. Subsequently, Section 5.3 reviews the most recent CM-DML decoding technique published in the literature. Furthermore, as the outcome of this chapter is two-fold, firstly Section 5.3.1 explores the benefits of using the proposed model in the reviewed CM techniques and, on the other hand, Section 5.5 develops a further investigation on novel constellations for CM-DML, along with simulations showing a substantial performance improvement with respect to the previously published work.

5.2 Frequency chirp in DMLs

Injecting a driving current signal, $I(t)$, into a DML produces an output optical signal which can be expressed as the electric field, $E(t)$, given by

$$E(t) = \sqrt{P(t)} \exp(j\varphi(t)) \tag{5.1}$$

where $P(t)$ is the instantaneous optical power, $j = \sqrt{-1}$ and $\varphi(t)$ is the time evolution of the signal phase which has three contributions: (i) the emitted optical frequency of the laser, (ii) the frequency chirp due to intensity modulation, and (iii) the phase noise. Under a complex baseband representation (i) is omitted, and $\varphi(t)$

is expressed in terms of (ii) and (iii) as

$$\varphi(t) = \varphi_{\Delta f}(t) + \varphi_{pn}(t) \tag{5.2}$$

where $\varphi_{\Delta f}(t)$ is related to the frequency chirp $\Delta f(t)$ as (see Equation (2.5) in page 37)

$$\begin{aligned}\varphi_{\Delta f}(t) &= 2\pi \int \Delta f(t) dt \\ &= 2\pi \int \frac{\alpha}{4\pi} \left(\frac{d}{dt} \ln P(t) + \kappa P(t) \right) dt \\ &= \frac{\alpha}{2} \left(\ln P(t) + \int \kappa P(t)\, dt \right)\end{aligned} \tag{5.3}$$

and $\varphi_{pn}(t)$ is the phase noise. Substituting Equation (5.3) into Equation (5.2), the electric field in Equation (5.1) can be rewritten as

$$E(t) = \sqrt{P(t)} \exp\left[i \left(\frac{\alpha}{2} \left(\ln P(t) + \int \kappa P(t)\, dt \right) + \varphi_{pn}(t) \right) \right] \tag{5.4}$$

The problem that Equation (5.4) poses is that the actual phase of the electric field depends on its own past values due to, in one hand, the term with the time integral of the power and, in the other hand, the nature of the phase noise term which, as seen in Chapter 4, is modelled as a Wiener-Lévy process, and recalling from Equations (4.12) and (4.13):

$$\begin{aligned}\varphi_{pn}(t) &= \int \phi(t)\, dt \\ \phi(t + T_s) - \phi(t) &\sim \mathcal{N}\left(0, \sigma_p^2\right) \\ \sigma_p^2 &= 2\pi \Delta\nu T_s\end{aligned}$$

where $\Delta\nu$ was the full-width half-maximum (FWHM) laser linewidth, and T_s the sampling period in seconds. Moreover, it must be noticed that Equation (5.4) is only an approximation since it assumes that α is a constant, while it actually depends on the emitted power [136]. However, this effect is out of the scope of this thesis and will be neglected.

5.3 A previously reported CM-DML technique

Starting back in 2015, a series of publications reported the first results on CM-DML both in simulation and experimental setup [137–140]. In these works, the authors adopted a generalised way to overcome the dependence on the signal history, that is, by considering only the phase difference between two consecutive symbols —what is known as *differential detection*— such that

$$\begin{aligned}\Delta\varphi(t) &= \varphi(t) - \varphi(t-T_s) \\ &= (\varphi_{chirp}(t) - \varphi_{chirp}(t-T_s)) + (\varphi_{pn}(t) - \varphi_{pn}(t-T_s)) \\ &= \frac{\alpha}{2}\left(\ln\frac{P(t)}{P(t-T_s)} + \kappa\int_{t-T_s}^{t} P(\tau)\,d\tau\right) + \Delta\varphi_{pn}(t)\end{aligned} \tag{5.5}$$

This way, the time integral needs to be performed only within a symbol period and the phase noise difference between two consecutive symbols, $\Delta\varphi_{pn}(t)$, could be neglected, hence

$$\Delta\varphi_{pn}(t) \approx 0. \tag{5.6}$$

The electric field can, therefore, be represented in a differential manner such that

$$E(t) = \sqrt{P(t)}\exp\left[i\left(\Delta\varphi(t) + \varphi(t-T_s)\right)\right]. \tag{5.7}$$

Notice that only the phase is treated differentially due to its cumulative nature, whilst the power is a memoryless signal and can be treated normally. Finally, the time integral between two consecutive sampling points is further simplified by considering the average power, reducing Equation (5.5) to

$$\Delta\varphi(t) = \frac{\alpha}{2}\left[\ln\left(\frac{P(t)}{P(t-T_s)}\right) + \kappa\frac{P(t)+P(t-T_s)}{2}T_s\right] \tag{5.8}$$

Notice that the simplification made in Equation (5.8) assumes a square waveform and a perfect sampling at the centre of each symbol period, where the laser is also assumed to have settled down and be free from transient ringing.

The next step consists of estimating the parameters of the frequency chirp. In [140] a experimental technique that consists of driving the laser with two training binary sequences ("‘LHLHLH...”’ and "‘LLHHLL...”’) is performed to, only at the

initial stage, obtain the following two auxiliary parameters

$$c_1 = \frac{\Delta\varphi_{L\to H} - \Delta\varphi_{H\to L}}{2\left(\ln P_H - \ln P_L\right)} \tag{5.9}$$

$$c_2 = \frac{\Delta\varphi_{H\to H} - \Delta\varphi_{L\to L}}{2\left(P_H - P_L\right)} \tag{5.10}$$

which are in turn related to the laser parameters α and κ as follows

$$c_1 = \frac{\alpha}{2} \tag{5.11}$$

$$c_2 = \frac{\alpha\kappa T_s}{4} \tag{5.12}$$

After detection at the receiver, with the recovered signal power $P(t)$ and differential phase $\Delta\varphi(t)$ it is possible to apply maximum-likelihood sequence estimation (MLSE) [141], which sequence length is reduced to two symbols due to the differential approach. A version of the Viterbi Algorithm (VA) [142] is proposed in [137–140]. The DSP procedure is detailed in Table 5.1 as explained in [139].

Table 5.1 – Viterbi Algorithm

Initialisation	
$t \leftarrow 1$	
$\hat{\mathbf{s}}(x_1) \leftarrow x_1$	
$\mathbf{\Gamma}(x_1) \leftarrow 0$	
Sequential operation	
1.	$\mathbf{\Gamma}(x_{t+1}, x_t) \leftarrow \mathbf{\Gamma}(x_t) + \lambda(\chi_t) \;\; \forall\chi_t = (x_{t+1}, x_t)$
2.	$x_t \leftarrow \arg\min_{x_t} \mathbf{\Gamma}(x_{t+1}, x_t)$
3.	$\hat{\mathbf{s}}(x_{t+1}) \leftarrow x_t$
4.	Update $\mathbf{\Gamma}(x_{t+1})$
5.	$t \leftarrow t + 1$

In this procedure, the VA algorithm is fed with the intensity and differential phase sequences, $P(t)$ and $\Delta\varphi(t)$, respectively. Two registers are defined: the sequence survivors are stored in $\hat{\mathbf{s}}(x_t)$ and the survivors sum of transition distances

are stored in $\mathbf{\Gamma}(x_t)$. Each state x_t is given by the power of the constellation point. The decoded sequence is the one at which the VA trace-back converges. There are M possible states $\{x_t\}$ and M^2 possible transitions $\chi_t \triangleq (x_t, x_{t-1})$. The probability of each transition $P(\chi_t)$ is simplified by means of the distance given by

$$\lambda(\chi_t) = |P(t-1) - x_{t-1}| + \left| \sqrt{P(t)} \cdot \exp[j\Delta\varphi(t)] - \sqrt{x_t} \cdot \exp[j\Delta\varphi_{E,t}] \right| \quad (5.13)$$

where $\Delta\varphi_{E,t}$ is the precomputed differential phase associated to the transition given by the pair of states (x_t, x_{t-1}) and Equation (5.8).

5.3.1 Proposed modification of the VA algorithm

In the derivation of the CM-DML approach it was assumed that the differential contribution of the phase noise was negligible. However, low-cost DFB lasers may have a non-negligible linewidth at baud rates around 10-20 GBaud. For instance, in [139,140] the employed DFB laser had a 10 MHz linewidth at the centre wavelength of 1550 nm, and it was modulated with a PAM4 signal at 10 Gbaud. Moreover, if there exists electrical noise in the driving signal $I(t)$, that will translate into not only a noisy $P(t)$ signal but also into noisy differential phase jumps $\Delta\varphi$.

Henceforth, it might seem sensible to modify the previously proposed VA algorithm in such a way that it incorporates a most robust metric in the presence of phase noise. It can be observed that the second term on the RHS of Equation (5.13) is an Euclidean distance used to (inversely) measure how likely it was a transition to have occurred. Here it is proposed to replace this distance by the proposed metric reported in Chapter 3, D^2_{prop}. Thus, the new computed transition distance will read as follows

$$\lambda'(\chi_t) = |P(t-1) - x_{t-1}| + D^2_{\text{prop}}\left(\sqrt{P(t)} \cdot \exp[j\Delta\varphi(t)], \sqrt{x_t} \cdot \exp[j\Delta\varphi_{E,t}] \right) \quad (5.14)$$

It has to be taken into account that the modelled D^2_{prop} was not defined for a differential detection scheme and therefore is not optimal. Nonetheless, it can still bring an improved phase noise sensitivity (as it will be seen in the next section). To that purpose, the AWGN parameter N_0 is input to the metric calculation in a reduced version as N_0/γ in order to enhance the phase noise variance σ_p^2. The

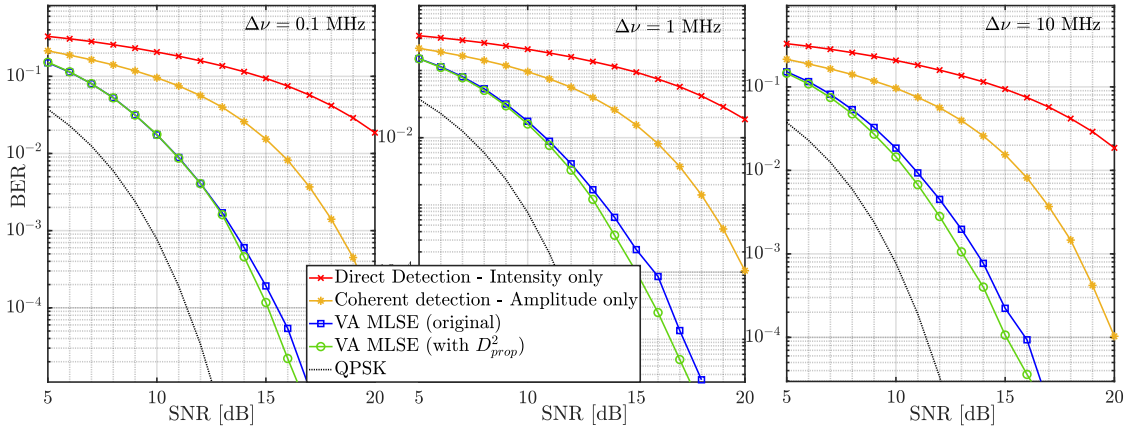


Figure 5.1 – Comparison of PAM4 with IMDD (crosses), Complex-amplitude detection (asterisks), original VA-MLSE (squares), and modified VA-MLSE (circles). Results shown for HD-BER versus SNR in dB for three values of $\Delta\nu$ = 100 kHz, 1 MHz, and 10 MHz.

factor γ will be empirically chosen and a further analysis is left for future research.

5.3.2 Simulation results

A set of numerical simulations have been carried out in order to evaluate whether the proposed metric can improve the performance of the original VA in the presence of non-negligible phase noise. The results are shown in Figure 5.1 where the curves represent the HD-BER versus SNR in dB for three values of combined linewidth $\Delta\nu$: 0.1 MHz (left), 1 MHz (centre), and 10 MHz (right). The tested algorithms are the original VA-MLSE described in [139] (blue squares) and its modification proposed above (green circles). Additionally, a pure IMDD decoding (red crosses), a complex-amplitude only decoding (yellow asterisks), and the theoretical AWGN limit for QPSK (black dotted line), are included for reference. For each data point a sequence of $5 \cdot 10^5$ PAM4 symbols have been generated. The factor γ is set to 128, 8, and 4 for the corresponding three tested values of $\Delta\nu$. As it can be observed, the proposed algorithm shows a clear benefit that is better for higher values phase noise, showing SNR penalties of around 0.4 and 0.6 dB at BER equal to 10^{-3} and of 1 dB at BER equal to 10^{-4} for the two highest values of phase noise.

5.4 Static analysis of the chirped electric field

The expression Equation (5.4) shows that although the electric field is a complex-valued signal where symbols can be potentially encoded in a 2-dimensional plane, there exists only one degree of freedom: the optical power $P(t)$ that, in turn, will be achieved by driving the laser with the only signal which waveform can be synthesised at the transmitter, the current, $I(t)$ (or, analogously, a voltage, depending on the device input type).

Consider the differential phase given Equation (5.4) and assume the phase noise is negligible ($\varphi_{pn} = 0$). It can be observed that it is composed by two terms, corresponding to the *transient* and *adiabatic* frequency chirp mechanisms [143, 144]

$$\Delta\varphi(t) = \underbrace{\Delta\varphi(t)_{\text{static}}}_{\text{transient chirp}} + \underbrace{\Delta\varphi(t)_{\text{dynamic}}}_{\text{adiabatic chirp}} \tag{5.15}$$

$$\Delta\varphi(t)_{\text{static}} = \frac{2}{\alpha}\ln\frac{P(t)}{P(t-T_s)} \tag{5.16}$$

$$\Delta\varphi(t)_{\text{dynamic}} = \frac{2}{\alpha}\kappa\int_{t-T_s}^{t} P(\tau)\,d\tau, \tag{5.17}$$

here denominated as “static” and “dynamic”, respectively.

The static contribution only depends on the logarithm of the ratio between the current and the previous symbol intensities, $P(t)$ and $P(t-T_s)$, whichever waveform takes the optical power during the transition with duration T_s. Therefore, in a (de)modulation context where the signal is sampled at 1 sps, the symbol period T_s becomes irrelevant, enabling the investigation of methods to design constellations that meet certain criteria. The goal is to build a set $\mathcal{S}$ that contains the intensity of each symbol, P_m, with $m = 0, \ldots, M-1$. Among these methods, three types are identified according to which information the receiver will have access to:

- **Intensity:** the (de)modulation technique corresponds to the conventional pulse amplitude modulation (PAM) and is a suboptimal approach because the receiver ignores the phase information. It has been previously covered in Section 2.1.4 and whenever included in simulation results will be for comparison purposes.

- **Phase only:** in this case, the constellation design method consists of searching for a set of intensity values aimed to achieve equally spaced angles. This approach is quite straightforward but suboptimal, as it ignores the amplitude information. A method of this kind is proposed later in Section 5.5.1.

- **Amplitude and phase:** using a combination of both amplitude and phase leads to a better constellation design. Here, the method consists of optimising the distribution of symbol intensities P_m so that the resulting complex points maximise a given figure of merit. This approach is described in Section 5.5.2.

On the other hand, the dynamic contribution results from an integration along a whole sampling period which yields it a baudrate- and waveform-dependent term as well as more vulnerable against timing errors or *jitter*. Initially, it is convenient to ignore this term in order to simplify the analysis and design of new CM-DML constellations. Therefore, for the rest of this chapter the dynamic term will be ignored for constellation design purposes, whilst in simulations it will be used (as in Equation (5.8)) in the synthesis of the transmitted waveform.

It will be seen in the following sections that the static analysis serves as a valid starting point, not only for the design of new constellations tailored to DML, but also for establishing their performance upper-bounds in a system based on coherent detection. In the promising future of CM-DML, a comprehensive analysis of the dynamic term and its impact performance will be necessary. However, that is out of scope of this thesis and consequently left open for further research.

5.5 The spirally-shaped electric field

The electric field that results from considering only the static contribution of Equation (5.5) in Equation (5.7), is given by the expression

$$E_{static,k} = \sqrt{P_k} \exp\left[i \left(\frac{\alpha}{2} \ln \frac{P_k}{P_{k-1}} + \varphi_{k-1} \right) \right], \tag{5.18}$$

where the variable t has been replaced by the symbol index k, without any loss of generality, thanks to the time-independence property. Equation (5.18) describes a

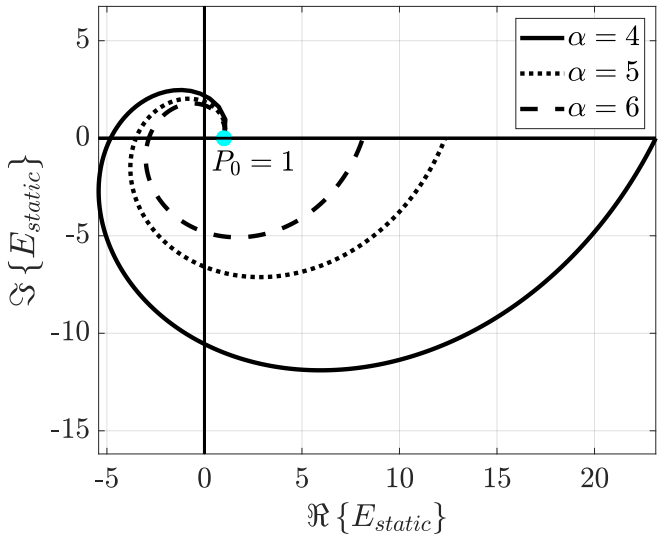


Figure 5.2 – Example spiral curves described by the complex E_{static} for three different values of α: 4 (solid), 5 (dotted), and 6 (dashed) using the static model of the chirp in Equation (5.18)

spiral curve that, starting at any point with polar coordinates ($\sqrt{P_0}$, φ_0), grows in amplitude while simultaneously rotates counter clock-wise. Figure 5.2 plots the evolution of three example E_{static} spirals for α equal to 4 (solid line), 5 (dotted line), and 6 (dashed line). The origin of the spirals is arbitrarily set to P_0 and $\varphi_0 = 0$ in all three cases.

An immediate consequence of the static analysis is that the spiral is scalable in its power, $|E_{static}|^2$. This is particularly useful while modelling a new constellation, since a first set of symbols can be obtained for a $P_0 = 1$ (simplifying further operations like $P_m/P_0 = P_m$), and posteriorly scale the set $\mathcal{S}$ to have a desired average transmitted power.

Apart from the initial coordinates (P_0, φ_0) and a final value P_f, a truncated spiral is entirely defined by the parameter α. In Figure 5.2 it can be observed that the smaller the α, the larger the required power excursion to produce the same phase shift. Within the framework of this chapter, unless stated otherwise, a value of $\alpha = 4$ is chosen to be used in both constellation design and system simulation, because it is a typical value for modern DFB semiconductor lasers and because it will allow a direct comparison with previously reported results in [139, 140]. A further analysis of the impact of α on performance limits is due for future investigation.

5.5.1 Spiral phase-shift keying

A simple way to start defining a possible spiral constellation is by only considering the phase jumps between symbols and set them to be equally spaced. This approach is named *spiral* phase shift keyed (SPSK) modulation and, although it is not optimum in terms of achievable capacity, the symbol powers are directly obtained by a simple calculation, as explained below.

The phase difference between any two consecutive symbols in a SPSK constellation, $\Delta\varphi_{SPSK}$, is obtained as

$$\Delta\varphi_{SPSK} = \frac{\lambda 2\pi}{M} \tag{5.19}$$

where M is the constellation order and λ is the *loop number*, that is, the number of 2π-loops that the spiral travels to accommodate the M constellation points. Since each symbol has a unique amplitude, λ can take integer values greater than 1, allowing two or more constellation points to share the same angular position.

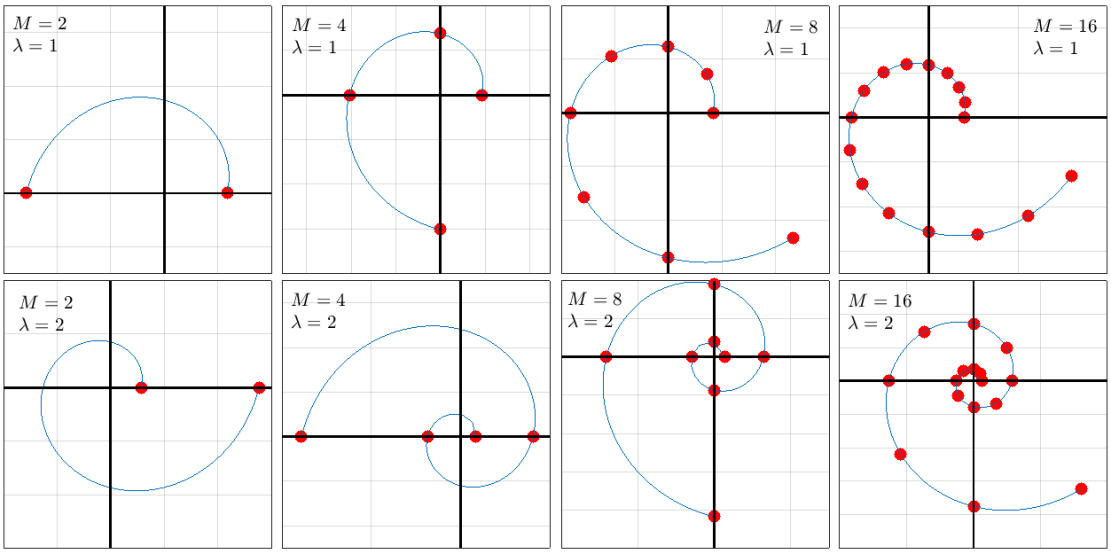


Figure 5.3 – Example of SPSK constellations for four values of M (2, 4, 8, and 16), two values of λ (1 and 2), and for $\alpha = 4$.

Equation (5.19) can be used with the static term in Equation (5.17) to define

$$\Delta\varphi_{SPSK,m} = \frac{m\lambda 2\pi}{M} = \frac{\alpha}{2}\ln\frac{P_m}{P_0}, \quad \text{with} \quad m \in \{0, 1, \ldots, M-1\} \tag{5.20}$$

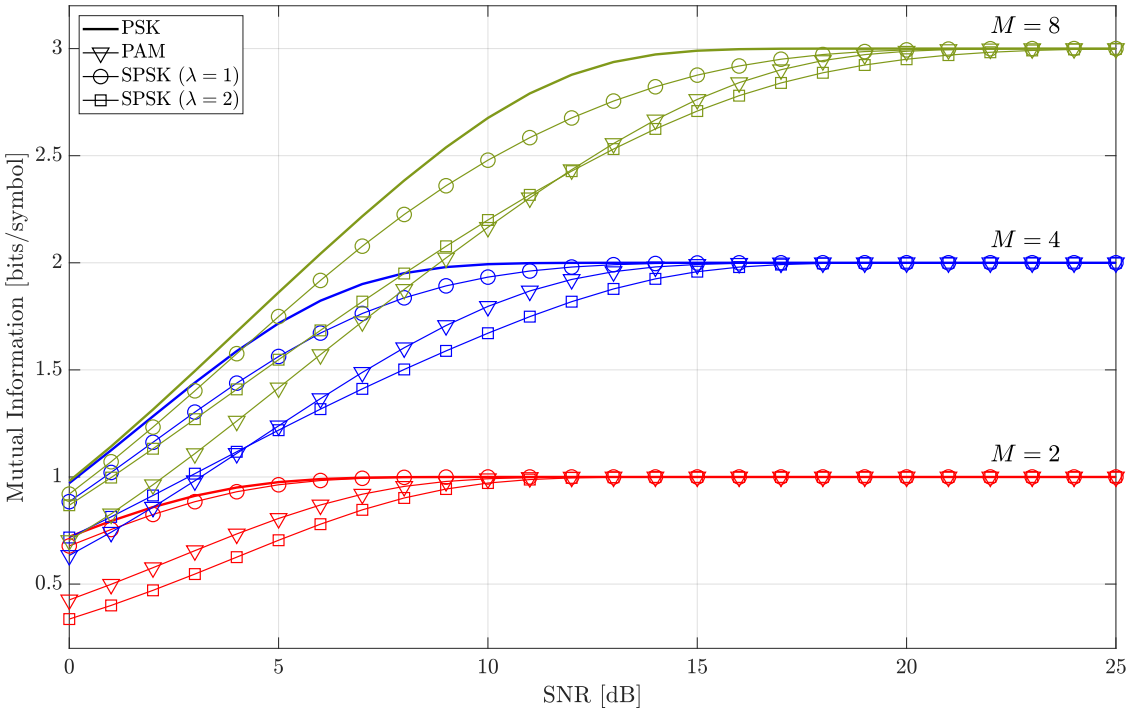


Figure 5.4 – AWGN-channel mutual information calculation for PSK (solid line), PAM (triangle markers), and SPSK constellations of orders M=2, 4, and 8 considering the static model of a directly modulated laser with $\alpha = 4$

where each P_m can be obtained as

$$\mathcal{S}_{SPSK} \triangleq \{P_m\} = \left\{ P_0 \exp\left(\frac{m\lambda 4\pi}{M\alpha}\right) \right\}, \qquad m = 0, \ldots, M-1 \tag{5.21}$$

expression that can be considered the generating function for the SPSK modulation format. A set of possible SPSK constellations is depicted in Figure 5.3 where some values of M (2, 4, 8, 16) and λ (1, 2) have been used.

In order to establish an upper-bound in the performance of the SPSK format, it is useful to compute the MI of the constellations for the AWGN channel. Figure 5.4 shows the obtained MI versus SNR for SPSK constellations with folding factor $\lambda = 1$ (circles) and $\lambda = 2$ (squares). For comparison, the MI for conventional PSK and PAM formats are also shown (solid line and triangles, respectively). Results are plotted for three values of M, namely 2 (red), 4 (blue), and 8 (green). The conclusion is that in the absence of phase noise, PSK presents the best performance in all cases, while SPSK with $\lambda = 1$ is only slightly worse, specially at intermediate

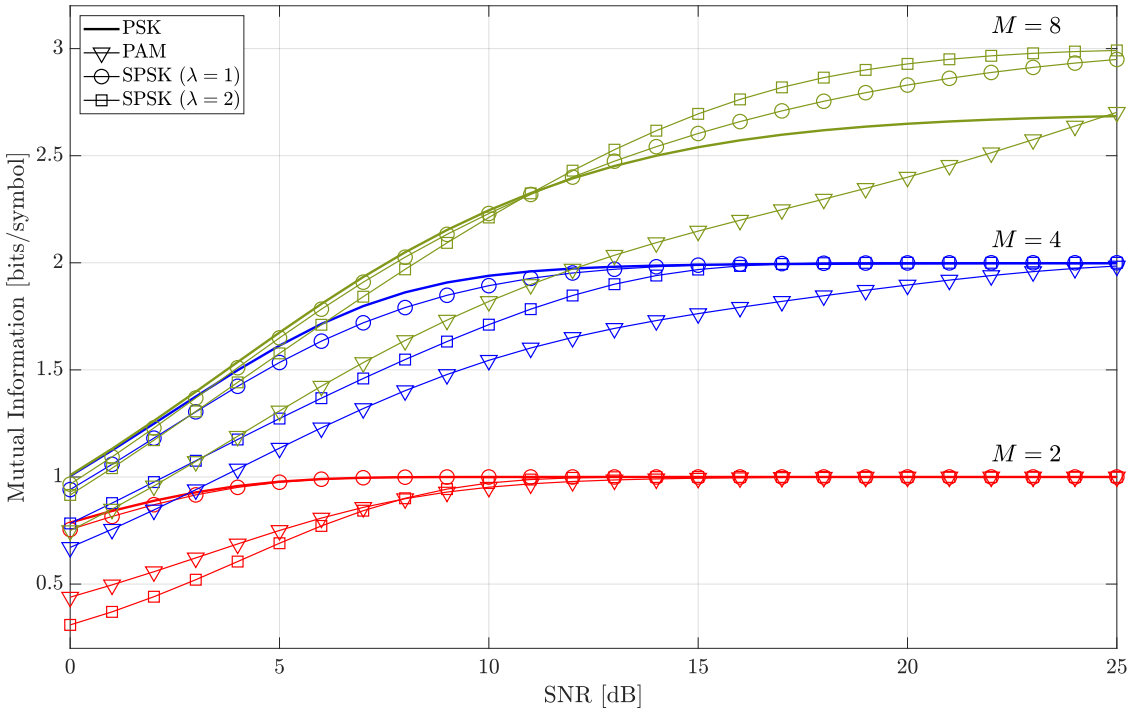


Figure 5.5 – PCAWGN-channel mutual information calculation for PSK (solid line), PAM (triangle markers), and SPSK constellations of orders M=2, 4, and 8 considering the static model of a directly modulated laser with $\alpha = 4$. Phase noise considered for a combined linewidth $\Delta\nu = 100$ MHz at a baudrate of 12.5 GBaud.

values of SNR and exhibiting a worst-case penalty of 0.05 bits per symbol with $M = 8$ and SNR = 11. Clearly, the conventional PAM and the SPSK with $\lambda = 2$ are worse candidates. For $M = 2$ and 4, the former is generally better than the latter although with $M = 8$ the situation is inverted for SNR < 11.

On the other hand, the situation changes when phase noise is introduced and the MI is computed according the PCAWGN channel model. As it can be seen in Figure 5.5, now SPSK with $\lambda = 2$ is only slightly worse than PAM for $M = 2$. Additionally, both versions of SPSK outperform the PSK format for severe values of phase noise ($\Delta\nu = 100$ MHz at 12.5 GBaud) with $M = 8$ and SNR is greater than 11 dB.

A discussion about the SPSK implications of the α factor on properties such as average power, extinction ratio, and modulation index is included in appendices C.1, C.2, and C.3, respectively.

5.5.2 Spiral amplitude and phase modulation

It was mentioned in Section 5.5.1 that the SPSK constellation was suboptimal in terms of achievable capacity, but convenient due to its simple implementation. This section describes an alternative and less straightforward way to build a spiral constellation, assuming a coherent receiver that has full access to the phase and amplitude information for demodulation. This new technique will be referred to as spiral amplitude and phase modulation, or SAPM.

The SAPM optimisation algorithm

The SAPM method consists of finding the distribution of M points along the spiral such that a chosen criterion is optimised. In this work, two criteria are contemplated: (i) the maximisation of the minimum distance (SAPM-D), and (ii) the maximisation of mutual information (SAPM-MI). Due to the lack of an analytical solution for either of the SAPM constellations, it is needed to rely on an iterative method. In this work, the SAPM optimisation algorithm is implemented in MATLAB® as described by the following pseudocode:

```
phi_0 <= 0
phis  <= 2*pi*m/M with m = 1,...,M-1
phis_est <= fminsearch(J(...), phis);
phis_est <= [phi_0, phis_est]
C <= sqrt(exp(phis*2/alpha)) * exp(sqrt(-1)*phis)
```

where `C` (or $\mathcal{C}$) is the obtained constellation of complex symbols. The minimisation algorithm is handled by the function `fminsearch`, which starts with an initial set of angles `phis` and converges to the final estimated set of angles, `phis_est` . The function `fminsearch` aims to minimise a target function, `J(...)`, which depending on the chosen criterion, is one of these two

$$J_{\mathrm{D}} = -\min_{i,j} \left(\left| C\left\{i\right\} - C\left\{j\right\} \right| \right), \quad \forall i,j = 0,1,...,M-1 \ \text{ and } \ i \neq j \tag{5.22}$$

$$J_{\mathrm{MI}} = -\mathrm{MI}\left(\mathcal{C}, N_0, \sigma_p^2\right), \tag{5.23}$$

where N_0 is the AWGN power and $\sigma_p^2 = 2\pi\Delta\nu T_s$.

2-SAPM

Consider the simplest case where $M = 2$. Using a normalised power of unity the 2-SAPM-MI and 2-SAPM-D algorithms are executed, setting a constraint by which the angular excursion between the first and last points in $\mathcal{S}$ has to be within a 2π-loop, that is, $\Delta\varphi_{0\to M-1} \leq 2\pi$. The MI of the resulting constellations is plotted in Figure 5.6 and compared with that of the BPSK, 2-SPSK, and PAM2 (in this simulation PAM2 is not equivalent to OOK, but the set $\mathcal{S} = \{1, 2\}$ normalised to unity power.

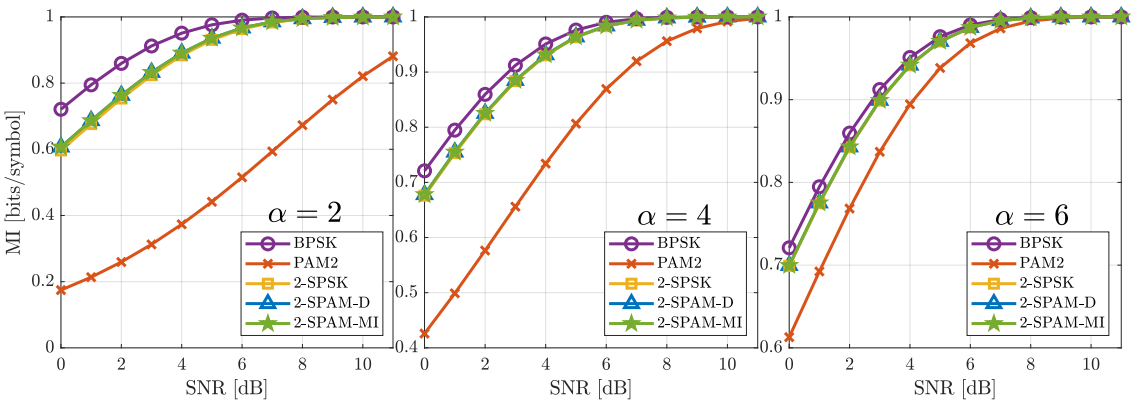


Figure 5.6 – Mutual information (MI) versus SNR computed for BPSK (circles), PAM2 (crosses), 2-SPSK (squares), 2-SAPM-D (triangles) and 2-SAPM-MI (stars) constellations. Three cases considered for α: from left to right, 2, 4, and 6. Phase noise is neglected ($\Delta\nu = 0$, AWGN model)

The results shown in Figure 5.6 indicate that any of spiral candidates results yields a higher MI than the one obtained by the PAM2. Another observation is that the spiral constellation (2-SPSK and 2-SAPM-D/MI) converge to almost identical conclusions, which makes sense because if there are only two points, maximising their MI is equivalent to maximise the distance between them. Finally, the spiral constellations suffer a penalty with respect to the best solution (BPSK) that decreases as the parameter α increases.

Notice that the PAM2 used in Figure 5.6 is suboptimal because their symbol powers have been arbitrarily chosen. The parameter α has an important impact on the distance between to points in a spiral. A discussion about the angle at which the distance between them is maximised is included in Appendix B.

Higher-order SAPM

In the context of complex modulation, the aim is to exploit the coherent detection capabilities in order to use higher order modulation formats. In this sense, 2-SPSK and 2-SAPM do not provide an improvement over the conventional OOK or BPSK in terms of bits per symbol (1). Therefore, the case for $M = 4$ and $M = 8$ (2 and 3 bits per symbol, respectively) are of special interest, since PAM4 was already reported to show incredibly good results in CM-DML and PAM8, on the other hand, is also receiving increasing interest, especially in high-speed short-reach scenarios such as data centre interconnection [86]. In what follows, it will be shown that both offer substantial room for improvement.

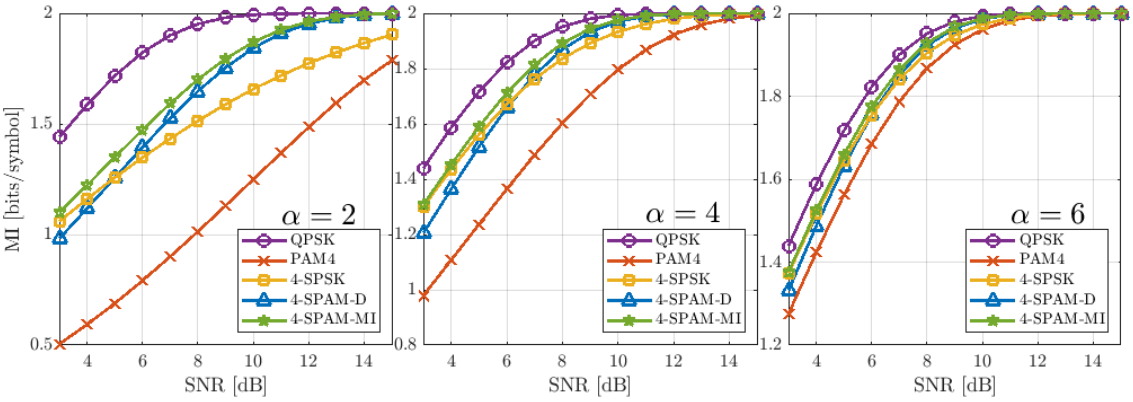


Figure 5.7 – Mutual information (MI) versus SNR computed for QPSK (circles), PAM4 (crosses), 4-SPSK (squares), 4-SAPM-D (triangles) and 4-SAPM-MI (stars) constellations. Three cases considered for α: from left to right, 2, 4, and 6. Phase noise is neglected ($\Delta\nu = 0$, AWGN model)

The SAPM algorithm is now performed with $M = 4$ using both criteria of maximising the minimum distance and the MI. During the iteration it has to be ensured that the constellation does not grow indefinitely so, before each call to the functions (5.22) or (5.23), the constellation is normalised to an average power of 1. The computed MI for different constellations is shown in Figure 5.7 for $M = 4$ and in Figure 5.8 for $M = 8$. In the first case QPSK is plotted as reference while for $M = 8$ the reference is set as 8-QAM. It can be seen that SAPM-MI outperforms the rest of formats in all cases. In general, all the spiral constellations show improvement over PAM, except for PAM8 with $\alpha = 6$ and mid values of SNR, where it becomes slightly better than the 8-SAPM-D (Figure 5.8, right plot).

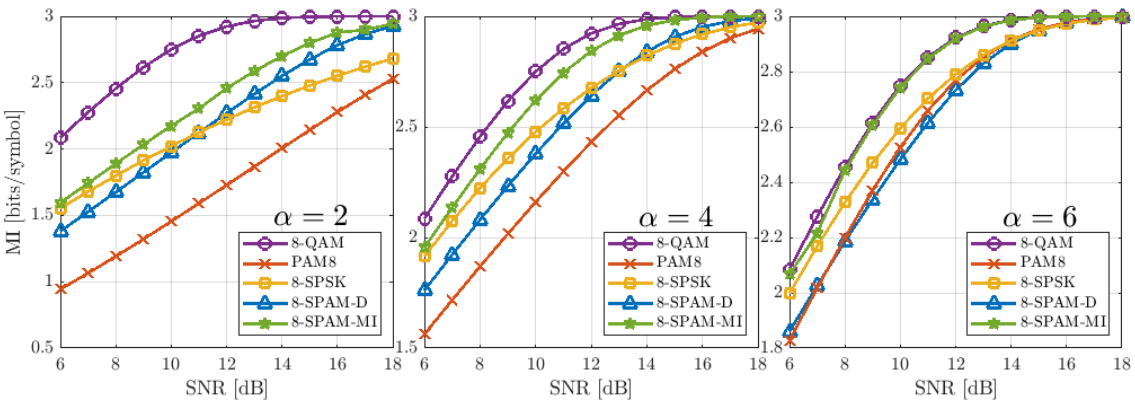


Figure 5.8 – Mutual information (MI) versus SNR computed for 8-QAM (circles), PAM8 (crosses), 8-SPSK (squares), 8-SAPM-D (triangles) and 8-SAPM-MI (stars) constellations. Three cases considered for α: from left to right, 2, 4, and 6. Phase noise is neglected ($\Delta\nu = 0$, AWGN model)

In terms of hardware complexity, the cost function Equation (5.22) is much more efficient than Equation (5.23), which means that the SAPM-D optimisation converges more rapidly than SAPM-MI. However, their solutions perform similarly well (particularly for the case $M = 4$) and that might be a good trade-off to go for an easier implementation.

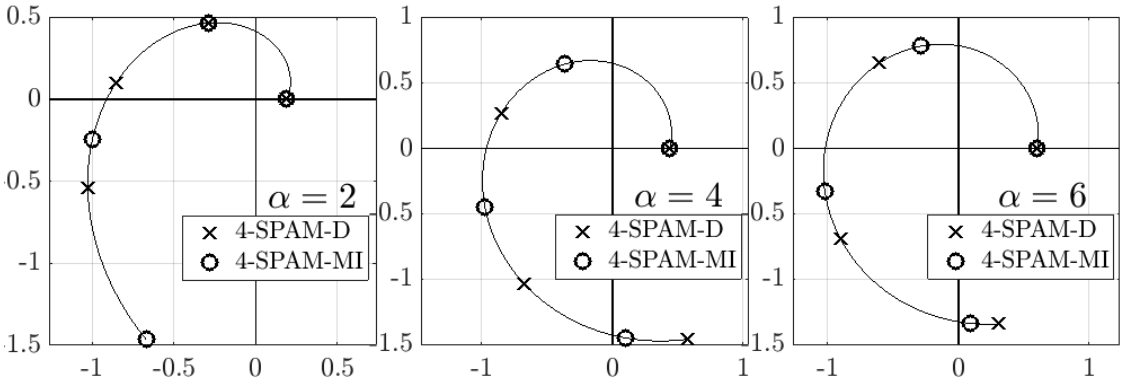


Figure 5.9 – Comparison of the 4-SAPM-D (crosses) and 4-SAPM-MI (circles) constellations for a given SNR = 6 dB, and for three values of α: 2 (left), 4 (centre), and 6 (right). In each plot, both constellations are normalised to have the same P_0 and lay on the same spiral.

For instance, to get an idea of how different 4-SAPM-D and 4-SAPM-MI look like, Figure 5.9 shows an example of their constellations superimposed, scaled such that they share the same P_0, thus laying on the same spiral curve. Notice, therefore,

that the two constellations shown in each plot do not have the same average power, but the MI is computed for a given SNR of 6 dB, so the results are equally valid. It is appreciated that for α equal to 2 and 4, the constellations find their place quite apart from each other along the spiral, but for $\alpha = 6$ (and higher) this distance decreases, as points tend to lay on the same ring for very high values of α.

Another important difference between SAPM-D and SAPM-MI is that the first has a unique shape for a given α, while the latter not only depends on α but also depends on the channel parameters, such that N_0 for the AWGN model or N_0 and σ_p^2 for the PCAWGN. Unless stated otherwise, assume $\sigma_p^2 = 0$. Figure 5.10 illustrates how the 4-SAPM-MI constellation evolves from low to high SNR for a fixed α, for which there are three cases considered: 2 (left), 4 (centre), and 6 (right). As expected, as SNR is swept, the points travel a longer distance along the spiral for lower values of α.

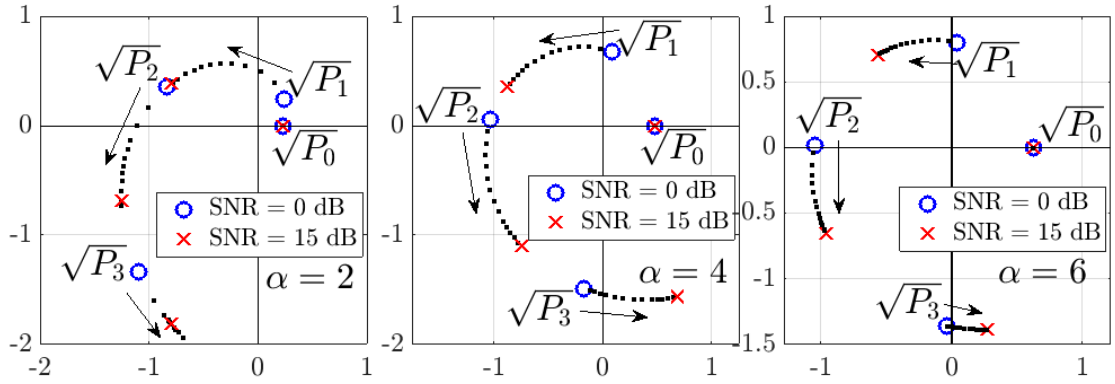


Figure 5.10 – Example of how the 4-SPAM-MI distribution of points evolves as the SNR is swept from 0 dB (blue circles) up to 15 dB (red crosses). The points have amplitudes $\sqrt{P_m}$ and the constellations are normalized to have the same P_0. Three values of α are shown, 2 (left), 4 (middle), and 6 (right).

To have a more insightful view, Figure 5.11 shows the probability densities for PAM4 (obtained from normalising the set $\mathcal{S} = \{1, 2, 3, 4\}$), 4-SPSK, and 4-SAPM-MI. From left to right, the used values of α are 3, 4, and 6, the SNR is set to 12 dB and a moderate amount of phase noise has been introduced ($\Delta\nu = 5$ MHz).

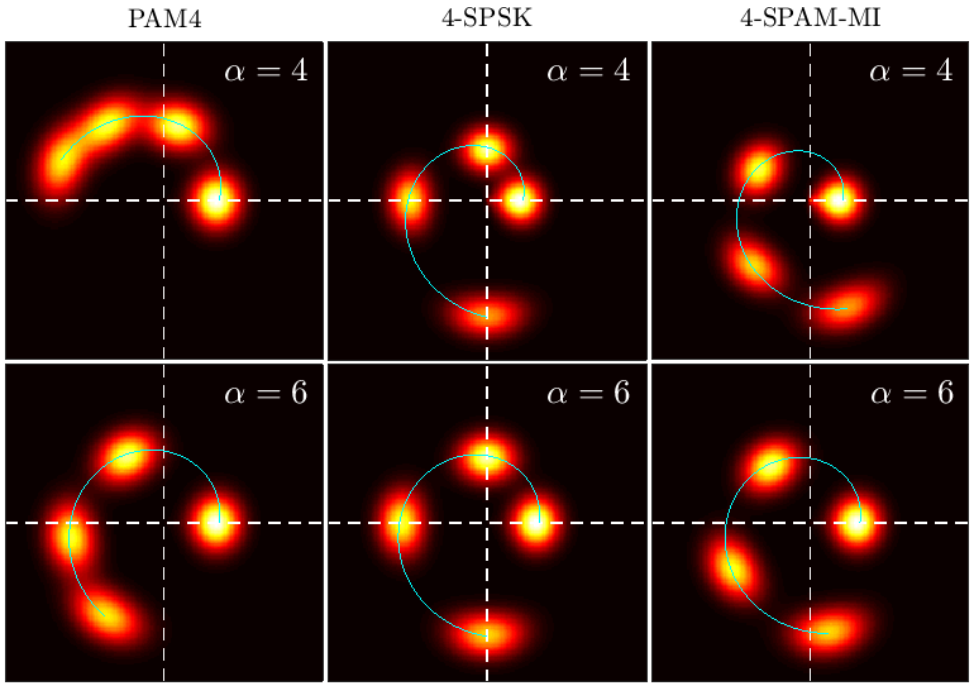


Figure 5.11 – Probability density of points in PAM4, 4-SPSK, and 4-SAPM-MI, using SNR = 12 dB and $\Delta\nu = 5$ MHz. Two cases for α: 4 (top) and 6 (bottom)

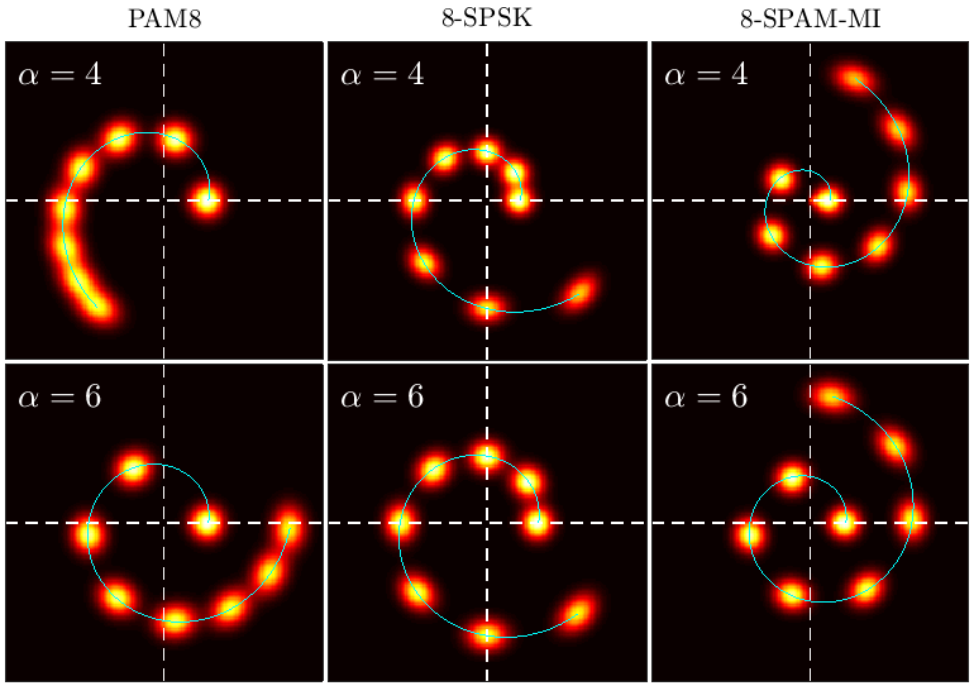


Figure 5.12 – Probability density of points in PAM8, 8-SPSK, and 8-SAPM-MI, using SNR = 16 dB and $\Delta\nu = 1$ MHz. Two cases for α: 4 (top) and 6 (bottom)

5.6 Simulation results

The previous sections have dealt with the design of constellations that, under the static model of the chirp, try to maximise criteria such as the minimum distance or the MI. The results obtained so far are promising and the question now is whether the presumed improvement offered by the spiral constellations do actually materialise after a system simulation.

In what follows, performance will be evaluated in terms of hard-decision SER (HD-SER) because Gray coding has not been defined for any of the spiral constellations and the comparison against a Gray-coded PAM would be unfair. The transmitted signal $x(k)$ is obtained from a random sequence of N symbols $D(k)$ that is used to select the power of the transmitted samples

$$|x(k)|^2 = \{P(k)\} = \{P(1), P(2), \ldots, P(N)\} \tag{5.24}$$

$$P(k) = \mathcal{S}(D(k)) \tag{5.25}$$

$$D(k) = U_k(\{0, \ldots, M-1\}) \tag{5.26}$$

where $U_k(\cdot)$ represents a function that returns a random object from the given set with uniform distribution, at the time instant $t = kT$, and recall that $\mathcal{S}$ is the set containing the constellation symbol powers. After the sequence of intensities is completed, the phase of $x(k)$ can be computed as

$$\varphi(k) = \Delta\varphi(k) + \Delta\varphi(k-1) \tag{5.27}$$

where the first symbol phase is computed using $\Delta\varphi(0) = 0$ and every update $\Delta\varphi(k)$ is obtained as in Equation (5.5) but using the simplification in Equation (5.8) and subtracting the average signal power from the (always positive) adiabatic chirp term as

$$\Delta\varphi(k) = \frac{\alpha}{2}\left[\ln\left(\frac{P(k)}{P(k-1)}\right) + \kappa\left(\frac{P(k)+P(k-1)}{2} - P_{avg}\right)T_s\right] \tag{5.28}$$

in order to take into account the frequency offset compensation. Lastly, the trans-

mitted signal is constructed as

$$x(k) = \sqrt{P(k)} \cdot \exp(j\Delta\varphi(k)) \tag{5.29}$$

to which there will be added AWGN and phase noise before being input to the VA algorithm for demodulation and error counting.

The first simulation experiment considers $M = 4$ for three values of α: 2, 4, and 6 ($c_1 = 1$, 2, and 3, respectively), and two values of phase noise are also considered, given by the combined linewidth $\Delta\nu$: 1 MHz and 10 MHz. The adiabatic coefficient c_2 is set to 0.5. For each data point, $5 \cdot 10^5$ are generated and the decoding algorithm is the VA-MLSE with the proposed modification described in Section 5.3.1.

Figure 5.13 plots the HD-SER versus SNR obtained with PAM4 (squares), 4-SPSK (asterisks), 4-SAPM-D (diamonds), and 4-SAPM-MI (stars). It can be seen that the spiral constellations, especially 4-SAPM-MI, perform better than PAM4 in most of the cases. The improvement is less significant for smallest value of α, being of the order of 0.1-0.2 dB, while for $\alpha = 4$ and $\alpha = 6$ up to 2 and 3 dB are observed. The only exception that can be appreciated is for high values of SNR with $\alpha \geq 4$ and $\Delta\nu = 10$ MHz, where both versions of 4-SAPM suffer a recession in their evolution, becoming the worst candidates. Although the reason for this has not been investigated in this work, it seems clear that the optimisation algorithms converge to a suboptimal local maximum. The SNR penalty of PAM4 versus 4-SAPM-MI for targeted HD-SER values of 10^{-2}, 10^{-2}, and 10^{-2} is summarized in Table 5.2.

Table 5.2 – SNR penalty for PAM4 versus 4-SAPM-MI at different HD-SER

M	$\Delta\nu$	α	HD-SER= 10^{-2}	HD-SER= 10^{-3}	HD-SER= 10^{-4}
4	1 MHz	2	0.5 dB	0.5 dB	0.2 dB
		4	1.4 dB	1.6 dB	2 dB
		6	2.2 dB	3 dB	3.4 dB
	10 MHz	2	0.3 dB	0.25 dB	0.25 dB
		4	0.7 dB	0.6 dB	0.1 dB
		6	1.3 dB	2.2 dB	2.6 dB

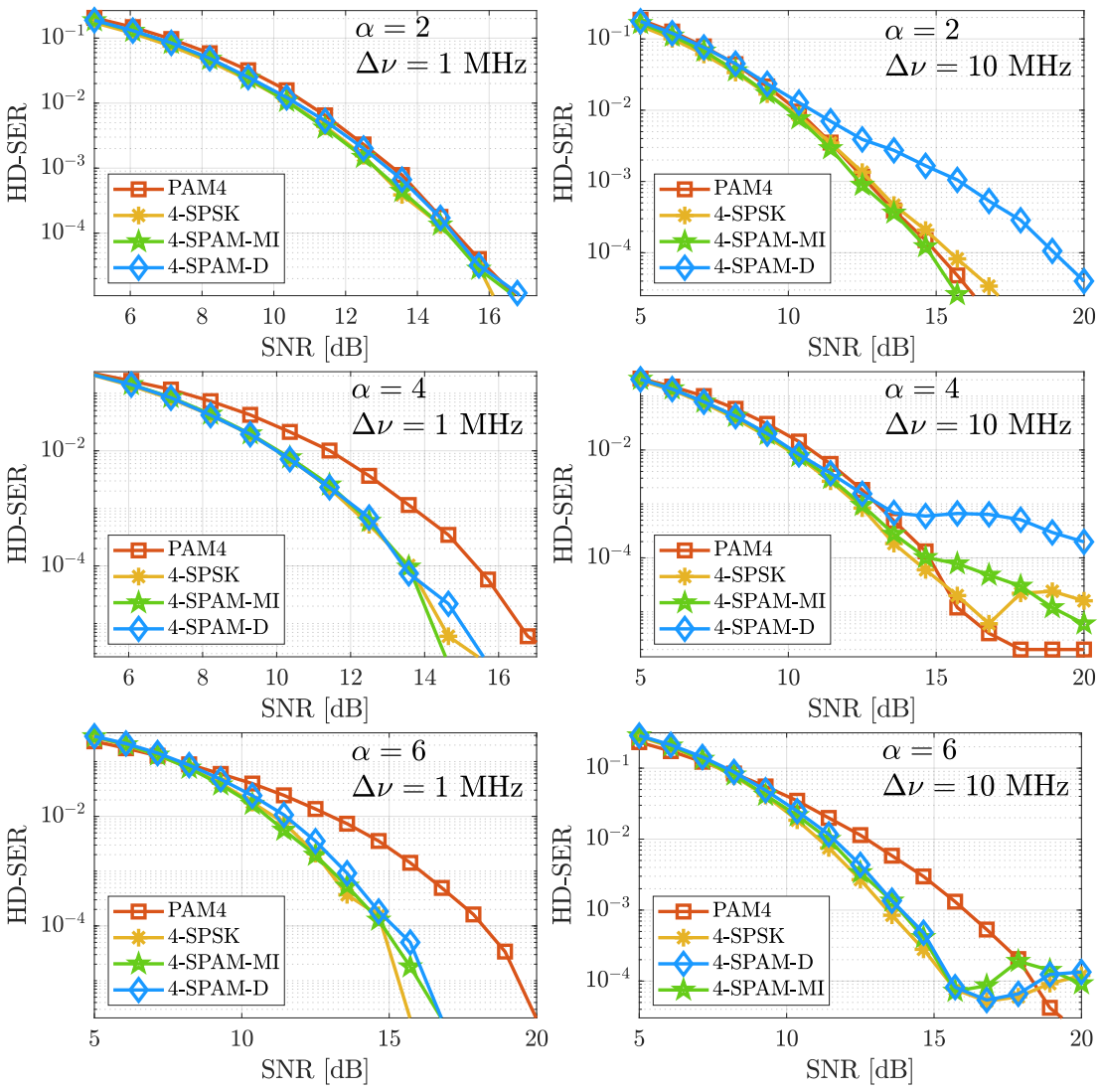


Figure 5.13 – Hard-decision symbol error rate (HD-SER) versus SNR for PAM, SPSK, and SAPM with $M = 4$- The parameter α is 2 (top), 4 (centre), and 6 (bottom). Phase noise given by $\Delta\nu = 1$ MHz (left) and 10 MHz (right). The adiabatic coefficient is set to $c_2 = 0.5$.

Table 5.3 – SNR penalty for PAM8 versus 8-SAPM-MI, at different HD-SER

M	$\Delta\nu$	α	HD-SER= 10^{-2}	HD-SER= 10^{-3}	HD-SER= 10^{-4}
8	1 MHz	2	0.7 dB	0.7 dB	1.6 dB
		4	3.8 dB	5 dB	5.1 dB
		6	6 dB	7.8 dB	8 dB
	10 MHz	2	1.7 dB	2.1 dB	2.1 dB
		4	-1.7 dB	0.5 dB	0.8 dB
		6	-3 dB	-3 dB	-3 dB

The next set of results is for constellations of order $M = 8$. The rest of simulation parameters are identical as in the previous case for $M = 4$, except for the SNR range that has been accordingly scaled up. In this case, as it can be seen in Figure 5.14, the spiral constellations perform significantly better than PAM8 with the lower phase noise ($\Delta\nu = 1$ MHz), where in the worst case for a HD-SER of 10^{-2}, the format 8-SAPM-MI exhibit a reduction in required SNR of 0.7, 3.8 and 6 dB for α values of 2, 4, and 6, respectively. The best case happens for a HD-SER=10^{-4}, where the improvements are of 1.6, 5.1, and 8 dB.

On the other hand, for the stronger phase noise ($\Delta\nu = 10$ MHz) the performance of the spirals is less consistent. With $\alpha = 2$, 8-SAPM-MI is clearly better than PAM8, reducing the SNR requirements by more than 1.7 dB, while for $\alpha = 4$ it exhibits an erratic behaviour alternating between a penalty of -1.7 dB with respect to PAM8 for HD-SER=10^{-2}, and an improvement of up to 0.8 dB for HD-SER=10^{-4}. Lastly, for $\alpha = 6$, 8-SAPM-MI sees its performance degraded for all the tested SNR values, showing a constant penalty of 3 dB for all three targeted HD-SER. Similarly, it is worth noting that the other two spirals, 8-SPSK and 8-SAPM-D, result in a very poor performance for the stronger phase noise scenario.

5.7 Conclusions

With the above results provided, it can be concluded that the static analysis of the frequency chirp in a DML can improve the existing approach of CM-DML.

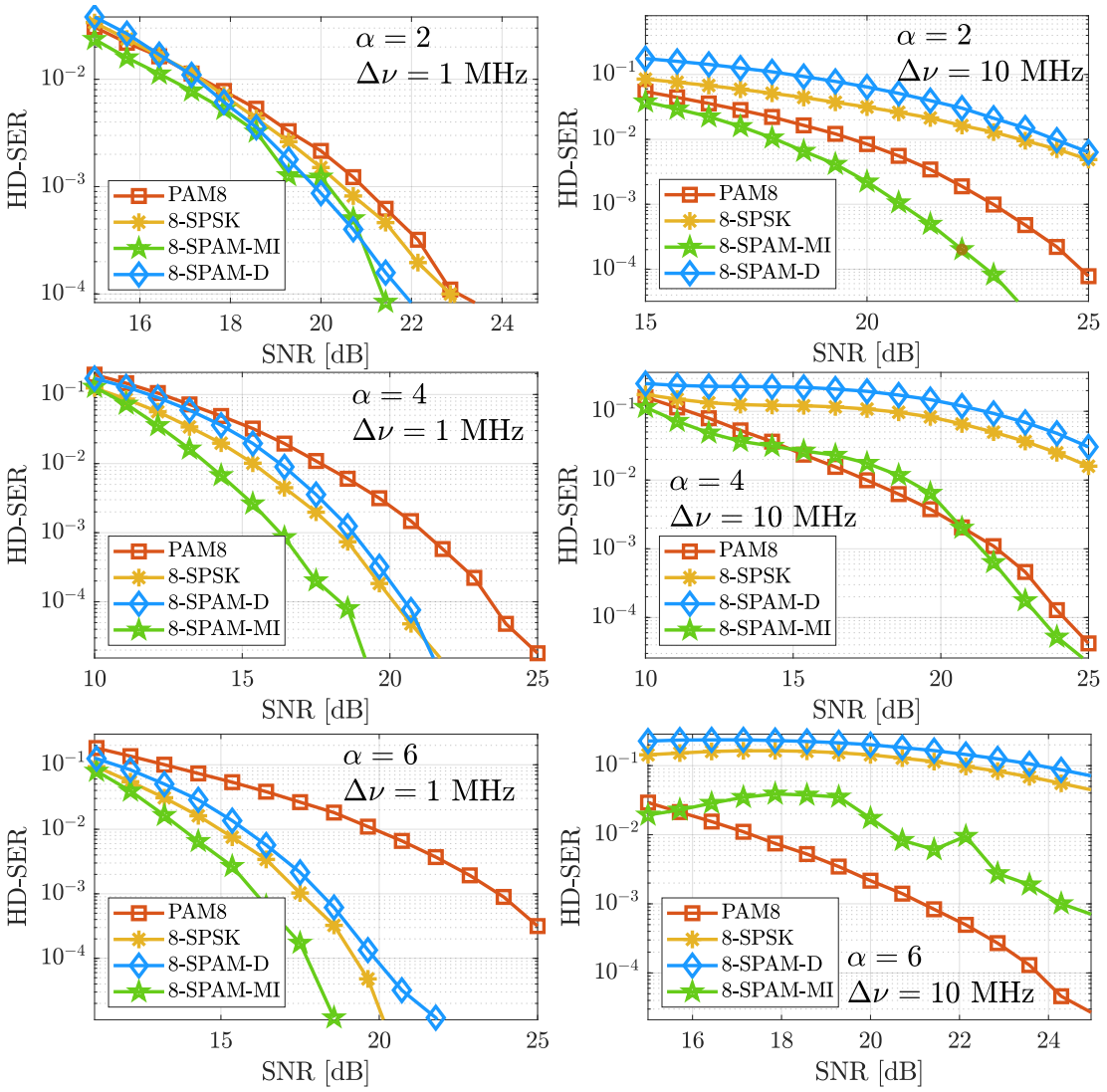


Figure 5.14 – Hard-decision symbol error rate (HD-SER) versus SNR for PAM, SPSK, and SAPM with $M = 8$- The parameter α is 2 (top), 4 (centre), and 6 (bottom). Phase noise given by $\Delta\nu = 1$ MHz (left) and 10 MHz (right). The adiabatic coefficient is set to $c_2 = 0.5$.

With a new set of optimised "spiral constellations" the VA-MLSE differential decoder can be made more robust to phase noise, particularly with the one that maximises MI (SAPM-MI) which has shown the best performance in the practical majority of numerical simulations performed. However, this technique is likely improvable, should a model of a differential channel be implemented. Chapter 6 suggests some ideas and preliminary results for a further research in this direction.

Chapter 6

Summary and future work

6.1 Summary

This thesis has dealt primarily with the problem of modelling the partially coherent AWGN channel (PCAWGN). It has been seen that this corresponds to a conventional AWGN channel with an added residual amount of phase noise. The PCAWGN channel usually arises in coherent communications systems where the signal phase, that continuously evolves with time, needs to be recovered by means of either hardware (PLL) or software (DSP) techniques. The AWGN present in the signal induces an error in the phase estimation process, yielding a spurious angular deviation that is commonly modelled with a Tikhonov distribution. Naturally, the additional source of noise decreases the link SNR and reduces the achievable capacity of the theoretical AWGN channel. This reduction is accentuated if the PCAWGN channel is treated with AWGN-like metrics, like the Euclidean distance.

This work highlights the importance of how a proper model of the PCAWGN channel can help mitigate this reduction in achievable capacity and the end-to-end benefits that can be obtained from it. Particularly, the focus of this thesis is centred in coherent optical communication channels. An introduction to the context of application in nowadays optical coherent networks, along with the challenges posed by the need of cost-sensitive solutions required in short-range links, is given in Chapter 1. The basic concepts of an optical communication system are reviewed in

Chapter 2, enumerating the main elements for transmission, transport and reception of optical signals, as well as the most popular modulation formats and techniques. Special attention has been paid to coherent detection architectures supported by subsequent DSP stages.

The optical PCAWGN channel has been covered in Chapter 3, where the mathematical model of the channel conditional probability has been derived, yielding a complex equation involving the sometimes untractable modified Bessel functions. As one main contribution of this work, an approximated model of the PCAWGN channel in polar coordinates has been proposed. It has been shown through simulations that the Gaussian approximations used in the proposed model are valid assumptions and that the reduced-complexity proposed approximation closely matches the exact solution when the channel MI is computed for QAM modulation formats of different sizes. The proposed approximation was also found to outperform the previously published Foschini's model, another approximation derived in cartesian coordinates by means of vector algebra.

Being the Euclidean distance so present in many DSP algorithms, and being clearly suboptimal in the presence of phase noise, it seemed mandatory to explore the possible benefits of the proposed model in practical applications, not only at the receiver side, but also at the transmitter. With this purpose, Chapter 4 collects some of the applications that have been investigated, such as in slicing circuits, carrier phase estimation or blind joint estimation of noise variances. With the simulations carried out it has been shown the robustness against phase noise of the proposed model, sensibly improving the Euclidean distance when both are tested to perform the same task.

Among the considered applications, that of complex modulation of semiconductor lasers, covered apart in Chapter 5, is of great interest. Its recent groundbreaking results shown in short-range systems using low-cost directly modulated lasers and coherent detection, make it a profitable field of study, with an immediate application in current, and next generation systems. In this topic, some methods to compute the amplitude levels for direct modulation have been described. These have been called the spiral constellations, because of the shape of the chirped electric field produced under direct intensity modulation of a laser. The new constellations have been compared against the conventional PAM formats, showing significant

improvements in most of the scenarios considered in simulations. Additionally, the CM-DML technique has been improved by upgrading the probabilistic metric in the decoding algorithm by that one studied in Chapter 3 for the PCAWGN channel, also exhibiting a higher tolerance to typical amounts of phase noise.

Lastly, the next section presents some other ideas that are regarded as the best guidelines to continue this research work in the future.

6.2 Future work

At the time of completion of this work, two main areas with potential for further research are identified. The first one concerns the approximated model of the PCAWGN channel and is discussed in Section 6.2.1. The second one, covered in Section 6.2.2, is related to the CM-DML and the spiral constellation subjects. In both cases, a primary idea is suggested, like the inclusion of nonlinear noise in the first case or the implementation of a differential channel model in the second case. The primary ideas are then followed by some other aspects that are not well developed at the time of writing this thesis, but they are considered equally important and potentially relevant in other than the already reviewed applications.

6.2.1 On the approximated PCAWGN channel model

Modelling nonlinear phase noise

An interesting way to continue the research presented on the PCAWGN channel would be trying to upgrade the proposed model to include the effects of nonlinear phase noise (NLPN) [145]. The sources of NLPN are diverse, well-known, and widely studied in the literature. To name some, consider the Kerr effect of the optical fibre [94], the Gordon-Mollenauer effect originated in amplified links [146], or the signal distortions caused by self-, and cross-phase modulation (SPM and XPM) [94, 147]. The impact of NLPN will therefore be more significant in longer distances, higher transmitted powers, and denser-channel systems. Although each source of NLPN has its own probability distribution, often non-Gaussian [66, 148], their phase contribution generally depends on the signal amplitude, $A(t)$, or power

$P(t)$. Consequently, the goal is to manipulate the proposed channel model to introduce an intensity-dependent phase term, $\varphi_{NL}(P(t))$, which will try to adapt the nonlinear channel and reflect phase deviations as NLPN would cause them.

As an initial approach, it can be considered that the NLPN contribution is directly proportional to the observed power, e.g. as

$$\varphi_{NL}(t) = \gamma_{NL} P(t) \tag{6.1}$$

where the scaling factor γ_{NL}, with units of radians per Watt, will be computed or empirically found for a given set of system parameters. Subtracting Equation (6.1) from the angular distance in Equation (4.6), a suggested "nonlinear" metric can be

$$D^2_{\text{prop}} = \Delta\rho^2 + \left(2\frac{N_0}{\sigma_p^2} + \frac{1}{|x|\,|y|}\right)^{-1} \cdot (\Delta\varphi - \varphi_{NL})^2 \tag{6.2}$$

Figures 6.1 and 6.2 show the probability density map and decision regions, respectively, using QAM constellations of orders 4, 8, 16, and 32 (from top to bottom), arbitrary SNR and nonlinear factors of 0.1, 0.3, and 0.5 rad/W (from left to right). The plots closely match previously published results (e.g. in [149]), motivating the belief that Equation (6.2) could be used to approximate NL-PCAWGN channels.

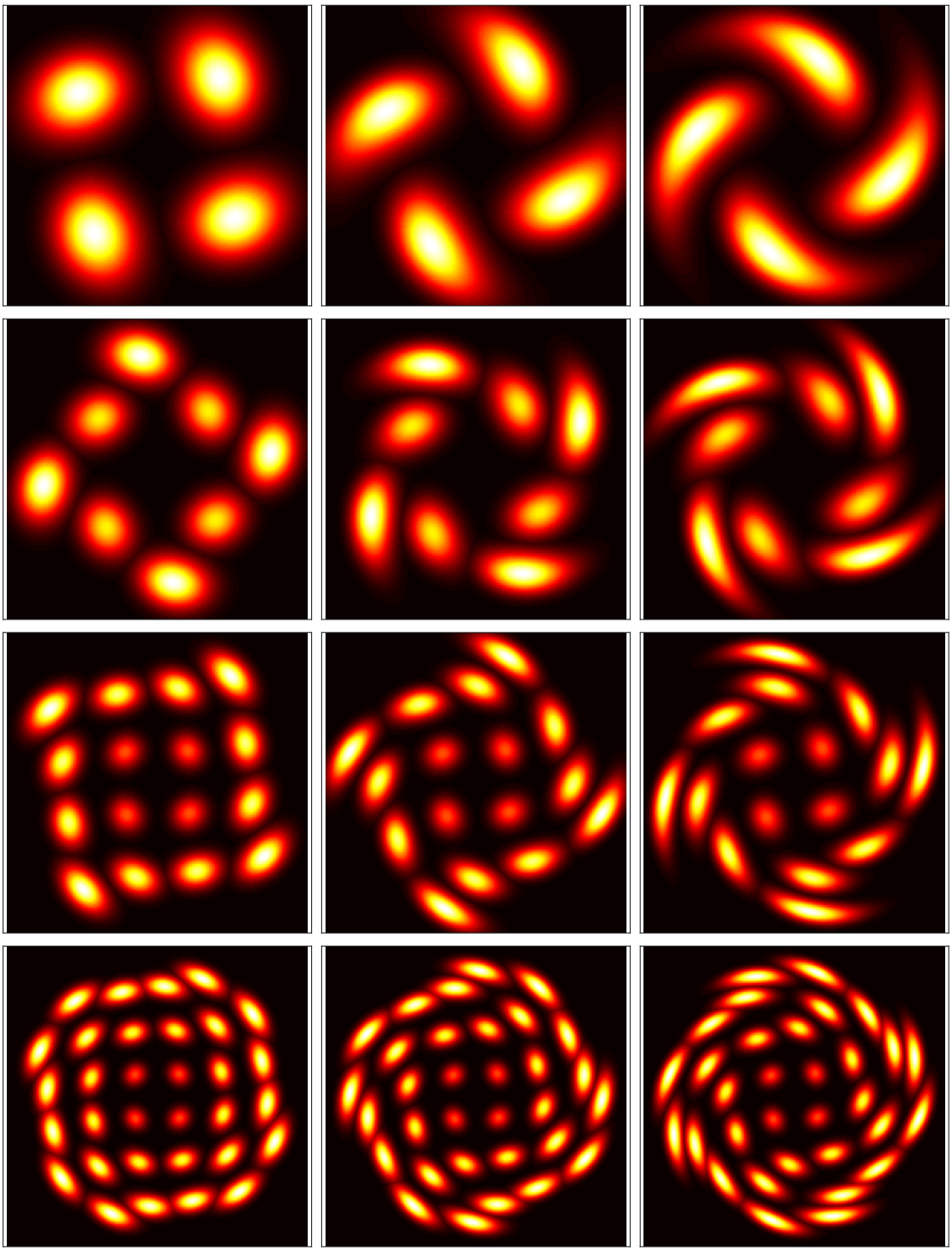


Figure 6.1 – Probability density obtained with the suggested D^2_{NL} for QAM of orders (and SNR), from top to bottom, 4 (8 dB), 8 (12 dB), 16 (16 dB), and 32 (20 dB). The nonlinear factor, γ_{NL} is set to, from left to right, 0.1, 0.3, and 0.5 rad/W.

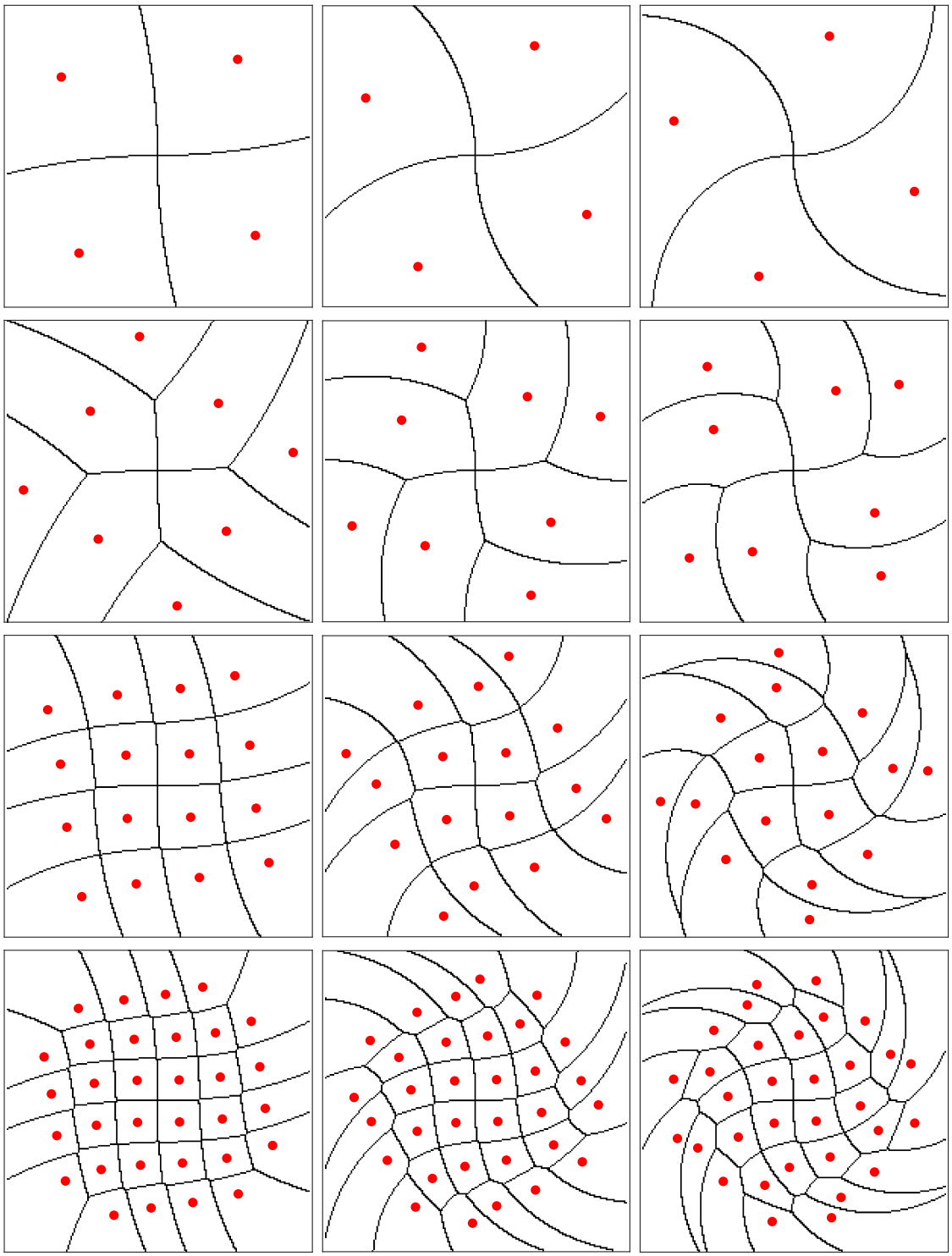


Figure 6.2 – Decision regions obtained with the suggested D^2_{NL} for QAM of orders (and SNR), from top to bottom, 4 (8 dB), 8 (12 dB), 16 (16 dB), and 32 (20 dB). The nonlinear factor, γ_{NL} is set to, from left to right, 0.1, 0.3, and 0.5 rad/W.

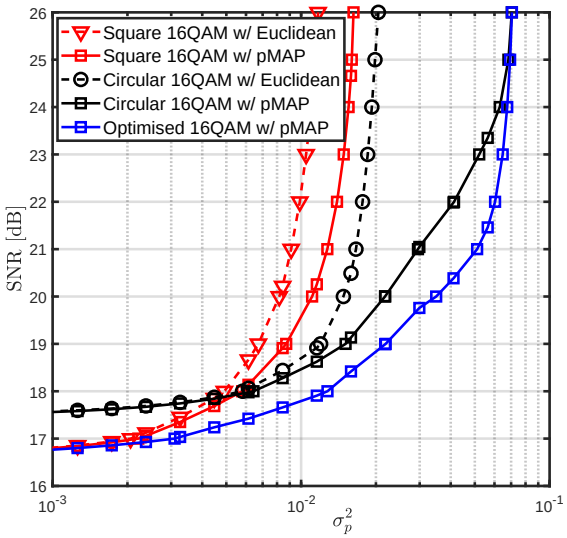


Figure 6.3 – Preliminary results for the suggested geometrical shaping optimisation algorithm for $M = 16$ (blue) versus square (red) and circular (black) 16-QAM. Curves show the contour for a target BER of $3.8{\cdot}10^{-3}$ for each pair of SNR and residual phase noise variance σ_p^2. The results with Euclidean slicers are shown for square (triangles) and circular (circles) QAM.

Constellation shaping

Another topic that has been initiated in this work, but left for future research, is *constellation shaping* (CS), a technique used to increase the MI of the channel by adapting the location of the symbols in the complex plane (*geometrical* shaping), or by adapting the probability of occurrence of each symbol (*probabilistic* shaping). CS requires a maximisation algorithm where the geometrical approach tries different locations and the probabilistic one does so with different distributions. In both cases, it is required to use a sort of metric or distance that stems from the channel model.

Numerous techniques can be found in the literature for both geometrical (e.g. in [150, 151]) and probabilistic shaping (e.g. in [152]). A more recent trend is the use of machine learning techniques for CS [153], which can become computationally expensive and prohibitive for cost-sensitive applications. This is why it is believed that the proposed approximated PCAWGN model can serve as a powerful tool for CS techniques.

As a very preliminary idea, the following simulation was performed: for each

given value of SNR and residual phase noise, σ_p^2, an optimisation algorithm returns the constellation that shows the lowest measured SER. The size of the constellation is $M = 16$ and the optimised constellation is constrained to be $\pi/2$-symmetrical. To further reduce the complexity, the algorithm assumes one of the 4 remaining points to be fixed, only solving the location of the other 3 points in a quadrant. Before computing the current BER, the algorithm normalises the constellation to a predefined reference power.

Figure 6.3 shows the results obtained from this simulation. The curves are plotted for the targeted BER $= 3.8 \cdot 10^{-3}$. The constellations considered are square and circular 16-QAM, and the optimised one. The results obtained using the Euclidean slicer are also shown for reference. The optimised constellation clearly outperforms the rest, converging to the circular QAM when $\sigma_p^2 \gg N_0$ and to the square QAM when $\sigma_p^2 \ll N_0$.

Channels with memory

Throughout this work, the channel under study has been always considered memoryless. That was so because it was assumed a prior signal equalisation that removed the memory introduced by the channel transfer function, and a later stage of CPE which removed the memory introduced by the the phase random evolution.

A reduced-complexity channel model with memory would help other techniques to save resources, or to use them for computing longer sequences. One of these techniques is MLSE [141], where a sequence of N symbols ($N-1$ transitions) has a probabilistic distance associated and is used to perform soft decisions. Similarly, other works study memory channels in order to find channel capacity boundaries [120, 154–156].

The proposed idea assumes that the channel only suffers from memory due to the random phase evolution and that a sequence of $N = 2n + 1$ symbols is fetched. Based on the concepts covered in Chapter 3, approximating in polar coordinates yields a memoryless radial dimension and the angular dimension contains all the memory of the channel. This can be further simplified because the angular distances have two components, AWGN-induced and phase noise, where only the latter introduces memory (since AWGN is non-cumulative). Hence, it can be noticed that

the reference sample at $t = 0$ will contain phase noise with variance σ_p^2. In a look-forward manner, the next samples will have a variance $2\sigma_p^2$, $3\sigma_p^2$, and so forth. This can be summarised by saying that for each possible sequence of N symbols, the suggested channel model can compute the probabilistic distance as

$$D\left(\chi_s\right) = \sum_{i=-n}^{n} \frac{\left(\rho\left(n\right) - \left|x_s\left(n\right)\right|\right)^2}{N_0} + \frac{\left(\arg\left(y\left(n\right)\right) - \arg\left(x_s\left(n\right)\right) - \varphi_s\left(n\right)\right)^2}{\left(\frac{N_0}{2\rho(n)|x(n)|} + \sigma_p^2\left(n\right)\right)} \tag{6.3}$$

where it must be noticed the dependence of σ_p^2 on the position of the sample within the sequence, n. It is well known that in a back-forward fashion, centred at sample $N + 1$, the best estimator is a Wiener filter, which obtains weighted taps that prioritise the central sample. The normalised variance $\sigma_p^2\left(n\right)/\sigma_p^2$ for each sample in a sequence of 5 symbols is shown in Figure 6.4 (black circles, left y-axis), while the weights here have been simply obtained as the inverse of the normalised variance for illustration purposes.

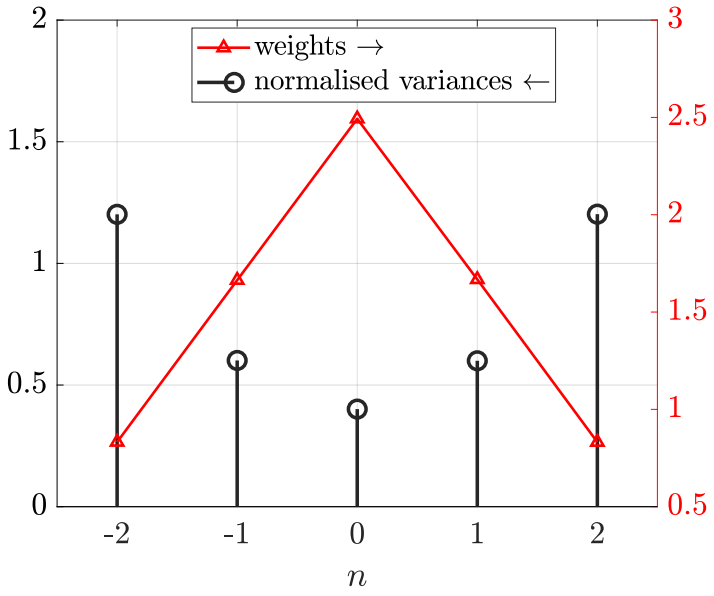


Figure 6.4 – Phase noise normalised variance (circles) of past and future samples with respect to the reference sample in the centre, exhibiting a shape like the well-known Wiener-filter tap coefficients inverted. Consequently, the weights (triangles) emphasise the central sample and decrease as it goes away from the centre.

A more detailed background on Wiener filters and its tap coefficient computation can be found in works like, e.g. [101, 102].

6.2.2 On the direct complex modulation of lasers

A differential channel model in polar coordinates

In Chapter 5 it was mentioned that optimising the constellation for a non-differential AWGN channel was a suboptimal solution and, as a consequence, the improvement predicted by obtained values of MI was not entirely consistent with what was observed in system simulations. A tentative workaround was scaling N_0 by a factor γ, empirically chosen, in order to enhance the phase noise against the AWGN. Then, it was suggested that a better solution would be to implement a differential channel model. At this point, this is an ongoing research from which some excellent results have been obtained so far.

It is convenient to start by defining what is called the *differential constellation*, $\mathcal{D}$, which contains $M \times M = M^2$ complex points $d_{1\to 2}$ and that can be represented in two ways:

$$\mathcal{D}_1 \triangleq \left\{ \sqrt{P_1} \exp\left(j\Delta\varphi_{1\to 2}\right) \right\} \tag{6.4}$$

$$\mathcal{D}_2 \triangleq \left\{ \sqrt{P_2} \exp\left(j\Delta\varphi_{1\to 2}\right) \right\} \tag{6.5}$$

where the difference is the power of which symbol is used to represent the modulus of the constellation point. The transition is defined to happen between symbol 1, or departing symbol, and symbol 2, or landing symbol[1]. As an example, consider an ideal PAM4 signal driving a laser with $\alpha = 4$, and neglect the adiabatic chirp, $c_2 = 0$. Figure 6.5 (a) shows the probability density of the electric field in the absence of phase noise (thus, looking static). This M points give raise to the $M^2 = 16$ points that represent the symbol transitions, which probability density is shown in Figure 6.5 (b) using the modulus of the departing symbol, $\sqrt{P_1}$, and (c) doing so with $\sqrt{P_2}$.

[1] this symbol numeration only has chronological sense, "from 1 to 2".

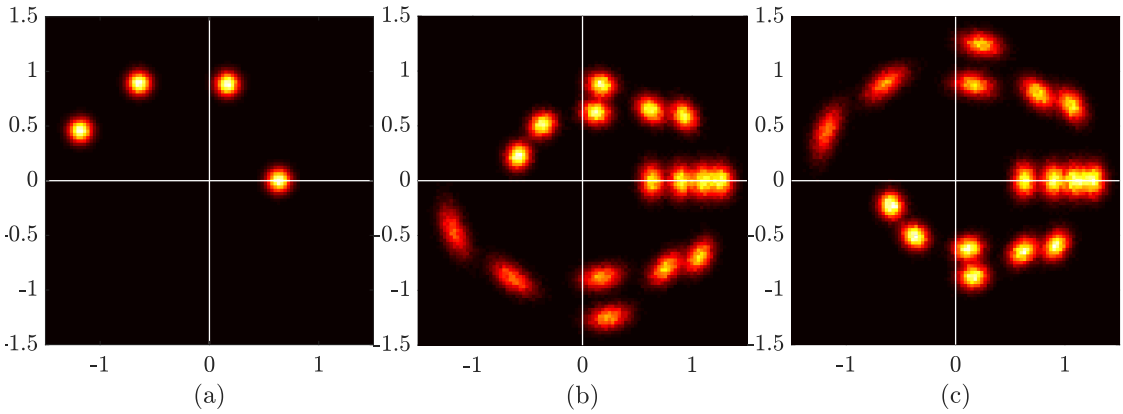


Figure 6.5 – Example of the probability density of the static spiral for PAM4 signal (a), and the differential constellation of transitions using $\sqrt{P_1}$ (b) and $\sqrt{P_2}$ (c).

Represented on the complex plane, $\mathcal{D}_1$ and $\mathcal{D}_2$ appear to be mirrored along the x-axis, and it might seem irrelevant which one is used but, as it will be seen, using $\mathcal{D}_2$ has a positive impact over using $\mathcal{D}_1$. Recall the probabilistic distance given in Equation (5.13) used to measure the transition probability in the VA algorithm, written as

$$\lambda\left(\chi_{1\to 2}\right) = \left|P_1 - x_1\right| + \left|\sqrt{P_2}\cdot\exp\left(j\Delta\varphi_{1\to 2}\right) - \sqrt{x_2}\cdot\exp\left(j\Delta\hat{\varphi}_{1\to 2}\right)\right|, \tag{6.6}$$

where the time index t has been replaced by the transition subscripts (1, 2), and which models a $\mathcal{D}_1$ differential constellation since it uses the symbol 1 in the transition, i.e. $P_1 \triangleq P\left(t-1\right)$. Figure 6.6 (a) shows the contour of equidistant curves obtained with Equation (6.6) superimposed to the probability density of the $\mathcal{D}_1$ constellation. It can be seen that not all the points are surrounded by the same number of contour lines, what might be interpreted as if some transitions are less likely than others.

A simple modification can be performed on the above distance, consisting of using $\mathcal{D}_2$ instead, as

$$\lambda'\left(\chi_{1\to 2}\right) = \left|P_2 - x_2\right| + \left|\sqrt{P_2}\cdot\exp\left(j\Delta\varphi_{1\to 2}\right) - \sqrt{x_2}\cdot\exp\left(j\Delta\hat{\varphi}_{1\to 2}\right)\right| \tag{6.7}$$

The resulting equidistant curves, superimposed to the $\mathcal{D}_2$ probability density, are shown in Figure 6.6 (b). Now the distribution seems to be better approximated,

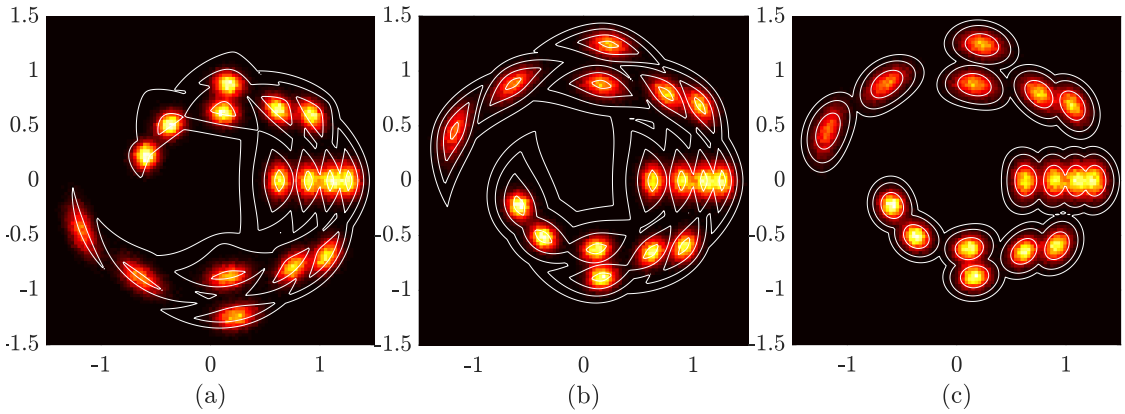


Figure 6.6 – Probability density of the 16 transitions for a PAM4 under complex modulation.

since all the 16 points are surrounded by the same number of contour lines and with the same distance values. It will be seen that this slightly modified metric yields an improvement over the original version.

The problem with these two versions is that none of them deals with the phase noise term nor the relatively higher phase shift that an AWGN corrupted point induces as it is closer to the origin. Hence, it is proposed to look into an alternative that exploits the properties of the PCAWGN approximated model presented in Chapter 3. The new distance will have the modulus distance squared as

$$D^2_{\text{modulus}} = \frac{\left|\sqrt{P_2} - \sqrt{x_2}\right|^2}{N_0} \tag{6.8}$$

and the observed differential angle $\Delta\varphi_{1\to2}$ is composed by four components: the expected phase jump $\Delta\hat{\varphi}_{1\to2}$ and three noise terms, the phase noise ($\Delta\varphi_{\text{pn}}$) and the AWGN-induced by the two symbols 1 and 2 ($\Delta\varphi_{\text{AWGN, 1}}$ and $\Delta\varphi_{\text{AWGN, 2}}$). The angular distance between observed and expected can be written as

$$\Delta^2\varphi_{1\to2} \triangleq \Delta\varphi_{1\to2} - \Delta\hat{\varphi}_{1\to2} = \Delta\varphi_{\text{AWGN, 1}} + \Delta\varphi_{\text{AWGN, 2}} + \Delta\varphi_{\text{pn}} \tag{6.9}$$

with $\Delta^2\varphi_{1\to2} \in [-\pi, \pi)$, and where the superscript in Δ^2 indicates the difference of a difference. As all the variables in Equation (6.9) are approximated as Gaussian, the observed angular distance distribution is also considered Gaussian, $\Delta^2\varphi_{1\to2} \sim \mathcal{N}\left(0, \sigma^2_{diff}\right)$ where the *differential* variance σ^2_{diff} is approximated as the sum of the

variances of the contributing terms, that is

$$\sigma^2_{diff} = \sigma^2_{\text{AWGN},1} + \sigma^2_{\text{AWGN},2} + \sigma^2_p \tag{6.10}$$

As it was seen in Section 4.1.1, the variance of the AWGN-induced noise is inversely proportional to an effective radius squared, ρ^2_{eff}, which gives a measure of how far is the point from the origin. In an *a posteriori* fashion, this radius was shown to be the geometrical mean between the transmitted and observed sample amplitudes. Thus, the AWGN-induced angular variances in the transition $1 \to 2$ are obtained as follows

$$\sigma^2_{\text{AWGN},1} = \frac{N_0}{2\sqrt{P_1 x_1}} \tag{6.11}$$

$$\sigma^2_{\text{AWGN},2} = \frac{N_0}{2\sqrt{P_2 x_2}} \tag{6.12}$$

Finally, inspired by the metric Equation (4.6) given in D^2_{prop}, a differential metric could take the following form

$$\delta D^2_{\text{prop}} = D^2_{\text{modulus}} + \frac{\left(\Delta^2\varphi_{1\to2}\right)^2}{2\left(\sigma^2_{\text{AWGN},1} + \sigma^2_p + \sigma^2_{\text{AWGN},2}\right)} \tag{6.13}$$

$$= \frac{\left|\sqrt{P_2} - \sqrt{x_2}\right|^2}{N_0} + \frac{\left(\Delta^2\varphi_{1\to2}\right)^2}{2\left(\sigma^2_{\text{AWGN},1} + \sigma^2_p + \sigma^2_{\text{AWGN},2}\right)} \tag{6.14}$$

$$= \frac{\left|\sqrt{P_2} - \sqrt{x_2}\right|^2}{N_0} + \frac{\left(\Delta^2\varphi_{1\to2}\right)^2}{2\left(\frac{N_0}{2\sqrt{P_1 x_1}} + \sigma^2_p + \frac{N_0}{2\sqrt{P_2 x_2}}\right)} \tag{6.15}$$

where the last expression can be simplified by multiplying it by N_0, yielding a reduced-complexity version expressed as

$$\overline{\delta D^2_{\text{prop}}} = N_0 \cdot \delta D^2_{\text{prop}} = \left|\sqrt{P_2} - \sqrt{x_2}\right|^2 + \frac{\left(\Delta^2\varphi_{1\to2}\right)^2}{\frac{1}{\sqrt{P_1 x_1}} + \eta + \frac{1}{\sqrt{P_2 x_2}}} \tag{6.16}$$

where η is the rotation factor described in Equation (4.7) and given by

$$\eta = 2\frac{N_0}{\sigma^2_p} \tag{6.17}$$

The equidistant curves obtained with the proposed metric are shown in Figure 6.6 (c) and they reflect an excellent match with the actual density distribution.

Initial results are now presented, where the system simulation parameters in [139] are replicated for a fairest comparison, and where the results for the previously published algorithm are plotted under the series name "Original", that is, using the transition distance in Equation (6.6). Similarly, the results obtained when using the transition distances given in Equations (6.7) and (6.16) are referred to as "Modified" and "Proposed", respectively. Figure 6.7 shows, on the top row, the computed BER after decoding the received signal with the VA-MLSE algorithm (as described in Section 5.3 and in [139]). The transmitted signal consists of a DML-chirped PAM4 with average power unity, with $5 \cdot 10^5$ symbols for each data point, the adiabatic coefficient is kept at $c_2 = 0.5$, the combined linewidth is set at $\Delta\nu = 10$ MHz, the baud rate is 12.5 Gbaud, and three α values are considered, 2 (left), 4 (centre), and 6 (right). The bottom row of Figure 6.7 shows the difference in required SNR, ΔSNR, versus the target BER, of the Proposed (solid line) and the Modified (dashed line) against the Original. Only the Modified distance can provide with SNR gains averaging to 0.4 dB, 0.7 dB, and 0.9 dB for the three values of α: 2, 4, and 6, respectively. More remarkable is the improvement observed with the Proposed metric, which averages to 1.2 dB, 1.75 dB, and 1.65 dB, respectively, for the same values of α.

These are excellent results for a preliminary evaluation, clearly indicating that this is a promising line of research. Probably the most natural step now would be using the Proposed distance and manipulate it to reflect probability rather than distance, and from there, try to use the spiral constellation optimisation algorithm (SAPM) in order to maximise the MI over a differential PCAWGN channel.

Power-constrained spirals

The methods described in Chapter 5 for the design of spiral constellations do not consider power limitations such as maximum optical power of the laser. What has been done so far is finding a set of intensity values that match a desired ratio between them, and scale it to have a desired average power. However, and particularly for lower values of α, a certain targeted ratio P_1/P_0 might require that $P_1 \gg P_0$. For

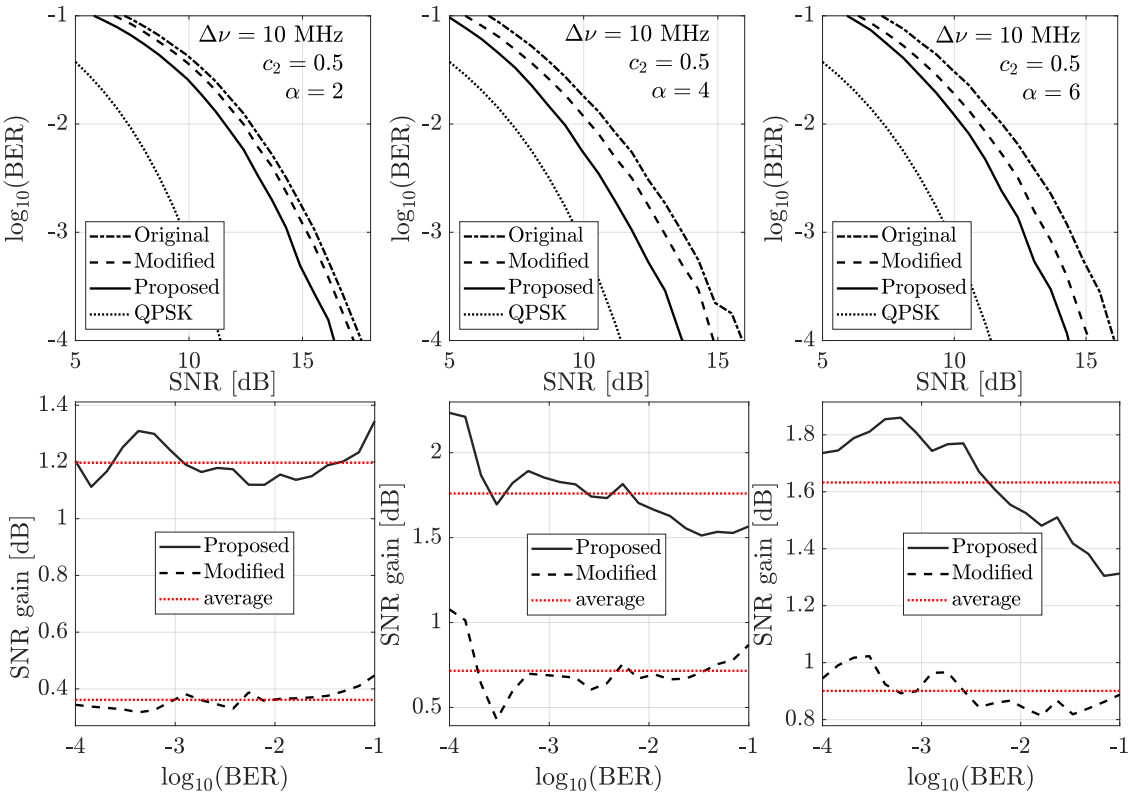


Figure 6.7 – CM-DML system simulation with the VA-MLSE decoding algorithm. (top) obtained BER versus SNR using the Original (dash-dotted), the Modified (dashed), and the Proposed (solid) transition distances. For each data point, a PAM4 signal with $5 \cdot 10^5$ symbols is generated. Three values of α are considered: 2 (left), 4 (centre), and 6 (right). (bottom) SNR gain versus targeted BER of the Proposed (solid) and Modified (dashed) metrics, along with the average for each case (red dotted).

instance, under the static model with $\alpha = 2$, a phase jump of $\Delta\varphi_{0\to1} = 3\pi/2$ radians requires a power ratio

$$\frac{P_1}{P_0} = \exp\left(\frac{\Delta\varphi_{0\to1}\alpha}{2}\right) = \exp\left(\frac{3\pi}{2}\right) \approx 111 \tag{6.18}$$

which means that modulating a 4-SPSK would require the highest power to be more than a hundred times the lowest one. To achieve this, either the high value can be pushed higher or the low value reduced. Both approaches pose a challenge, though, because the laser ideally has to be kept working in its linear region, above threshold and below the maximum allowed optical power.

Therefore, the realisability of a certain order SPSK constellation will be determined by the α parameter and the minimum/maximum optical power of the laser. Regarding the SAPM constellation, the optimisation algorithm can be modified so that the power values that limit the linear region of the laser are fixed. An added benefit of this approach is that only $M-2$ points have to be iteratively computed, versus the $M-1$ points of the original version.

It could be interesting to assess the impact of the dependence of α on the laser intensity [136] in the design of spiral constellations and whether a proper model of this phenomenon can provide any benefit to the decoding VA algorithm. It can also be questioned the feasibility of deliberately targeting a desired value of α during the device manufacturing process.

Experimental verification

Naturally, the ultimate goal is to perform an experimental verification of the proposed concepts for CM-DML. A standard testbed transmitter will consist of a high-speed arbitrary waveform generator (AWG) driving a DFB laser packaged and interfaced to support both DC bias and direct RF modulation. As a first approach, a back-to-back (B2B) connection to the receiver would suffice to assess the potential of the proposed constellation design methods and decoding metric. A crucial step is to find the α and κ parameters of the laser. A simple method to obtain an estimation of these parameters was proposed in [139], relying on a pair of basic training sequences of binary symbols. However, to eventually develop more accurate models, it might be convenient to characterise the laser with a higher number of parameters that appear in the laser rate equations. For such purpose, a popular method for the extraction of laser parameters based on the idea of frequency response subtraction is described in [157, 158] and similar ones can be also found in [159, 160]. An alternative method based on the measurements of RIN and phase noise is also proposed in [161].

Beyond the static approach where the laser is driven with a signal of M fixed amplitude levels, with a given set of laser parameters it could be possible to simulate the rate equations [162–166], permitting to shape a driving waveform that accounts for the adiabatic chirp contribution to the phase evolution. This would yield a

dynamic modulation technique where the amplitude levels for each symbol depend on the past transmitted signal. A waveform shaping technique tailored to DFB lasers is presented in [167], cancelling relaxation oscillations and reducing the laser dynamic memory.

References

[1] D. J. Richardson, "Filling the light pipe," *Science*, vol. 330, no. 6002, pp. 327–328, 2010.

[2] A. D. Ellis, J. Zhao, and D. Cotter, "Approaching the non-linear shannon limit," *Journal of Lightwave Technology*, vol. 28, pp. 423–433, 2 2010.

[3] J. T. Barnett, "Cisco Visual Networking Index VNI: Complete Forecast Update, 2017-2022," Dec. 2018. Accessed February 2020.

[4] A. Chralyvy, "Plenary paper: The coming capacity crunch," in *2009 35th European Conference on Optical Communication*, pp. 1–1, 9 2009.

[5] H. Waldman, "The impending optical network capacity crunch," in *2018 SBFoton International Optics and Photonics Conference (SBFoton IOPC)*, pp. 1–4, Oct. 2018.

[6] J. v. Weerdenburg, R. Ryf, J. C. Alvarado-Zacarias, R. A. Alvarez-Aguirre, N. K. Fontaine, H. Chen, R. Amezcua-Correa, T. Koonen, and C. Okonkwo, "138 Tbit/s Transmission over 650 km Graded-Index 6-Mode Fiber," in *2017 European Conference on Optical Communication (ECOC)*, pp. 1–3, Sept. 2017.

[7] G. Rademacher, R. S. Luis, B. J. Puttnam, T. A. Eriksson, E. Agrell, R. Maruyama, K. Aikawa, H. Furukawa, Y. Awaji, and N. Wada, "159 Tbit/s C-L Band Transmission over 1045 km 3-Mode Graded-Index Few-Mode Fiber," in *2018 Optical Fiber Communications Conference and Exposition (OFC)*, pp. 1–3, Mar. 2018.

[8] B. J. Puttnam, R. S. Luís, W. Klaus, J. Sakaguchi, J.-D. Mendinueta, Y. Awaji, N. Wada, Y. Tamura, T. Hayashi, M. Hirano, and J. Marciante, "2.15 Pb/s transmission using a 22 core homogeneous single-mode multi-core fiber and wideband optical comb," in *2015 European Conference on Optical Communication (ECOC)*, pp. 1–3, Sept. 2015.

[9] D. Soma, Y. Wakayama, S. Beppu, S. Sumita, T. Tsuritani, T. Hayashi, T. Nagashima, M. Suzuki, H. Takahashi, K. Igarashi, I. Morita, and M. Suzuki, "10.16 Peta-bit/s Dense SDM/WDM transmission over Low-DMD 6-Mode 19-Core Fibre Across C+L Band," in *2017 European Conference on Optical Communication (ECOC)*, pp. 1–3, Sept. 2017.

[10] K. Shibahara, T. Mizuno, D. Lee, and Y. Miyamoto, "Advanced MIMO signal processing for dense SDM transmission using multi-core few-mode fibers," in *2016 21st OptoElectronics and Communications Conference (OECC) held jointly with 2016 International Conference on Photonics in Switching (PS)*, pp. 1–3, July 2016.

[11] Y. Jung, Y. Jung, S. Jain, S. Alam, and D. Richardson, "Fibre-Based Components for SDM Systems," in *2017 Asia Communications and Photonics Conference (ACP)*, pp. 1–1, Nov. 2017.

[12] M. Jinno, "Spatial channel network (SCN): Opportunities and challenges of introducing spatial bypass toward the massive SDM era [invited]," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 11, pp. 1–14, Mar. 2019.

[13] C. F. Lam, *Passive optical networks: principles and practice*. Amsterdam; Boston: Elsevier/Academic Press, 2007. OCLC: 228148439.

[14] L. G. Kazovsky, W. T. Shaw, D. Gutierrez, N. Cheng, and S. W. Wong, "Next-Generation Optical Access Networks," *Journal of Lightwave Technology*, vol. 25, pp. 3428–3442, 11 2007.

[15] H. Rohde, S. Smolorz, E. Gottwald, and K. Kloppe, "Next generation optical access: 1 Gbit/s for everyone," in *2009 35th European Conference on Optical Communication*, pp. 1–3, Sept. 2009.

[16] F. O. Amaya Fernández, A. C. Soto, and I. Tafur Monroy, "Optimizing the next generation optical access networks," *IEEE Latin America Transactions*, pp. 438–443, 2009.

[17] Dayou Qian, N. Cvijetic, Junqiang Hu, and Ting Wang, "A Novel OFDMA-PON Architecture With Source-Free ONUs for Next-Generation Optical Access Networks," *IEEE Photonics Technology Letters*, vol. 21, pp. 1265–1267, Sept. 2009.

[18] M. Ruffini, N. Doran, M. Achouche, N. Parsons, T. Pfeiffer, X. Yin, H. Rohde, M. Schiano, P. Ossieur, B. O'Sullivan, and others, "DISCUS: End-to-end network design for ubiquitous high speed broadband services," in *Transparent Optical Networks (ICTON), 2013 15th International Conference on*, pp. 1–5, IEEE, 2013.

[19] K. Tanaka and A. Agata, "Next-generation Optical Access Networks for C-RAN," in *Optical Fiber Communication Conference (2015), paper Tu2E.1*, p. Tu2E.1, Optical Society of America, Mar. 2015.

[20] G. Vall-llosera, A. Rafel, N. Parkin, M. Angelou, D. Klonidis, I. Cano, M. Presi, G. Papastergiou, I. Tomkos, J. Prat, and E. Ciaramella, "COCONUT cost, power consumption and migration analysis: A route towards NG-PON3," in *2015 17th International Conference on Transparent Optical Networks (ICTON)*, pp. 1–4, July 2015.

[21] R. Sanchez, J. A. Hernandez, J. M. Garcia, and D. Larrabeiti, "Provisioning 1 Gb/s symmetrical services with next-generation passive optical network technologies," *IEEE Communications Magazine*, vol. 54, pp. 72–77, Feb. 2016.

[22] A. Davis, M. Pettitt, J. King, and S. Wright, "Phase diversity techniques for coherent optical receivers," *Journal of Lightwave Technology*, vol. 5, pp. 561–572, Apr. 1987.

[23] R. A. Linke, "Optical heterodyne communications systems," *IEEE Communications Magazine*, vol. 27, pp. 36–41, 10 1989.

[24] L. G. Kazovsky, "Phase- and polarization-diversity coherent optical techniques," *Journal of Lightwave Technology*, vol. 7, pp. 279–292, 2 1989.

[25] M. Taylor, "Coherent Detection Method Using DSP for Demodulation of Signal and Subsequent Equalization of Propagation Impairments," *IEEE Photonics Technology Letters*, vol. 16, pp. 674–676, Feb. 2004.

[26] S. J. Savory, "Digital filters for coherent optical receivers," *Opt. Express*, vol. 16, pp. 804–817, 1 2008.

[27] M. Kuschnerov, F. Hauske, K. Piyawanno, B. Spinnler, M. Alfiad, A. Napoli, and B. Lankl, "DSP for Coherent Single-Carrier Receivers," *Journal of Lightwave Technology*, vol. 27, pp. 3614–3622, Aug. 2009.

[28] H. Louchet, K. Kuzmin, and A. Richter, "Improved DSP algorithms for coherent 16-QAM transmission," in *2008 34th European Conference on Optical Communication*, pp. 1–2, Sept. 2008.

[29] X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, and G. Li, "Electronic post-compensation of WDM transmission impairments using coherent detection and digital signal processing," *Optics Express*, vol. 16, no. 2, p. 880, 2008.

[30] E. J. Bachus, R. P. Braun, W. Eutin, E. Großmann, H. Foisel, K. Heimes, and B. Strebel, "Coherent optical-fibre subscriber line," *Electronics Letters*, vol. 21, pp. 1203–1205, Dec. 1985.

[31] E. J. Bachus, R. P. Braun, C. Caspar, E. Grossmann, H. Foisel, K. Heimes, H. Lamping, B. Strebel, and F. J. Westphal, "Ten-channel coherent optical fibre transmission," *Electronics Letters*, vol. 22, pp. 1002–1003, Sept. 1986.

[32] P. Cochrane and M. Brain, "Eradicating the central office with coherent optical technology," in *1988 International Zurich Seminar on Digital Communications - Mapping New Applications onto New Technologies. Proceedings*, pp. 59–62, Mar. 1988.

[33] G. D. Cormack and J. R. Waschuk, "1995 coherent broadband access systems (optical fiber networks)," in *Conference Proceeding IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, pp. 431–434, June 1989.

[34] S. J. Savory, "Digital coherent optical access networks," in *2013 IEEE Photonics Conference*, pp. 125–126, Sept. 2013.

[35] Y. Yoshida, A. Maruta, and K. i. Kitayama, "Coherent IFDMA-PON: A novel green and elastic optical access networks," in *2012 IEEE Photonics Society Summer Topical Meeting Series*, pp. 155–156, July 2012.

[36] R. Gaudino, V. Curri, G. Bosco, G. Rizzelli, A. Nespola, D. Zeolla, S. Straullu, S. Capriata, and P. Solina, "On the use of DFB Lasers for Coherent PON," in *Optical Fiber Communication Conference (2012), paper OTh4G.1*, p. OTh4G.1, Optical Society of America, Mar. 2012.

[37] J. D. Reis, D. M. Neves, and A. L. Teixeira, "Transmission aspects on broadband coherent optical access networks," in *2011 IEEE Third Latin-American Conference on Communications*, pp. 1–6, Oct. 2011.

[38] H. Rohde, S. Smolorz, J. S. Wey, and E. Gottwald, "Coherent optical access networks," in *2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference*, pp. 1–3, Mar. 2011.

[39] A. Theurer, R. Seidel, R. Ziegler, C. Zawadzki, Z. Zhang, N. Keil, A. Matiss, and A. G. Steffan, "Polymer based hybrid integrated coherent receiver for next generation optical access networks," in *2012 Asia Communications and Photonics Conference (ACP)*, pp. 1–3, Nov. 2012.

[40] J. Berger, G. Sarlet, M. Stiller, and M. Weber, "400ZR specification update." `https://www.ofcconference.org/en-us/home/exhibit-hall/show-floor-programs/400zr-specification-update/`, Mar. 2019. The Optical Networking and Communication Conference & Exhibition.

[41] I. Lyubomirsky, B. Zhang, M. Sluyski, R. Baca, M. Filer, G. Nicholl, M. Nowell, and J. DeAndrea, "Toward baseline for 400GBASE-ZR optical specs." `http://www.ieee802.org/3/ct/public/adhoc/19_0221/lyubomirsky_3ct_01_190221.pdf`, Feb. 2019. IEEE P802.3ct Ad-Hoc Meeting.

[42] S. Sugimoto, K. Minemura, K. Kobayashi, M. Seki, M. Shikada, A. Ueki, T. Yanase, and T. Miki, "High-speed digital-signal transmission experiments by optical wavelength-division multiplexing," *Electronics Letters*, vol. 13, pp. 680–682, Oct. 1977.

[43] T. Miki and H. Ishio, "Viabilities of the Wavelength-Division-Multiplexing Transmission System Over an Optical Fiber Cable," *IEEE Transactions on Communications*, vol. 26, pp. 1082–1087, July 1978.

[44] W. J. Tomlinson, "Wavelength multiplexing in multimode optical fibers," *Applied Optics*, vol. 16, pp. 2180–2194, Aug. 1977.

[45] ITU-T, "G.694.2: Spectral grids for WDM applications: CWDM wavelength grid," 2003. Accessed August 2019.

[46] ITU-T, "G.694.1: Spectral grids for WDM applications: DWDM frequency grid," 2012. Accessed August 2019.

[47] F. Cavaliere, F. Ponzini, M. Presi, and E. Ciaramella, "Migration towards high speed optical access enabled by WDM techniques," in *2009 Asia Communications and Photonics conference and Exhibition (ACP)*, vol. 2009-Supplement, pp. 1–7, Nov. 2009.

[48] Y. C. Chung, "Recent advancement in WDM PON technology," in *2011 37th European Conference and Exhibition on Optical Communication*, pp. 1–3, Sept. 2011.

[49] J. D. Reis and A. L. Teixeira, "Architectural optimization of coherent ultra-dense WDM based optical access networks," in *2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference*, pp. 1–3, Mar. 2011.

[50] H. Mu, H. K. Shim, H. G. Choi, U. H. Hong, and Y. C. Chung, "A proposal for cost-effective 10-Gb/s, 10-GHz spaced ultra-dense WDM PON," in *2013 18th OptoElectronics and Communications Conference held jointly with 2013 International Conference on Photonics in Switching (OECC/PS)*, pp. 1–2, June 2013.

[51] J. Prat, M. Angelou, C. Kazmierski, R. Pous, M. Presi, A. Rafel, G. Vall-llosera, I. Tomkos, and E. Ciaramella, "Towards ultra-dense wavelength-to-the-user: The approach of the COCONUT project," in *2013 15th International Conference on Transparent Optical Networks (ICTON)*, pp. 1–4, June 2013.

[52] N. H. Zhu, Z. Shi, Z. K. Zhang, Y. M. Zhang, C. W. Zou, Z. P. Zhao, Y. Liu, W. Li, and M. Li, "Directly Modulated Semiconductor Lasers," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 24, pp. 1–19, Jan. 2018.

[53] M. Shirasaki, H. Nishimoto, T. Okiyama, and T. Touge, "Fibre transmission properties of optical pulses produced through direct phase modulation of DFB laser diode," *Electronics Letters*, vol. 24, pp. 486–488, Apr. 1988.

[54] R. S. Vodhanel, "5 Gbit/s direct optical DPSK modulation of a 1530-nm DFB laser," *IEEE Photonics Technology Letters*, vol. 1, pp. 218–220, Aug. 1989.

[55] I. N. Cano, A. Lerin, V. Polo, and J. Prat, "Direct Phase Modulation DFBs for Cost-Effective ONU Transmitter in udWDM PONs," *IEEE Photonics Technology Letters*, vol. 26, pp. 973–975, May 2014.

[56] A. Lerín, G. Chu, V. Polo, I. N. Cano, and J. Prat, "Chip integrated DFB-EAM for directly phase modulation performance improvement in UDWDM-PON," in *2015 European Conference on Optical Communication (ECOC)*, pp. 1–3, Sept. 2015.

[57] M. Presi, G. Cossu, R. Corsini, F. Bottoni, and E. Ciaramella, "Low cost coherent receivers for UD-WDM NRZ systems in access networks," in *16th International Conference on Transparent Optical Networks, ICTON*, IEEE Computer Society, 2014.

[58] M. Presi, F. Bottoni, R. Corsini, G. Cossu, and E. Ciaramella, "All DFB-Based Coherent UDWDM PON With 6.25 GHz Spacing and a >40 dB power Budget," *IEEE Photonics Technology Letters*, vol. 26, pp. 107–110, Jan. 2014.

[59] M. Presi, R. Corsini, M. Artiglia, and E. Ciaramella, "Using directly modulated DFBs without power penalty in low-cost and high-power budget coherent access networks," *Electronics Letters*, vol. 50, Mar. 2014.

[60] I. N. Cano, A. Lerín, G. Y. Chu, V. Polo, and J. Prat, "Performance comparison between direct phase modulated DFB and RSOA for cost effective transmitter in udWDM-PONs," in *2015 17th International Conference on Transparent Optical Networks (ICTON)*, pp. 1–4, July 2015.

[61] I. N. Cano, J. C. Velásquez, and J. Prat, "7.5 Gb/s direct DFB phase modulation with 8-DPSK for 6.25 GHz spaced coherent UDWDM-PONs," in *2016 Optical Fiber Communications Conference and Exhibition (OFC)*, pp. 1–3, Mar. 2016.

[62] I. N. Cano, A. Lerin, and J. Prat, "DQPSK Directly Phase Modulated DFB for Flexible Coherent UDWDM-PONs," *IEEE Photonics Technology Letters*, vol. 28, pp. 35–38, Jan. 2016.

[63] D. Kim, B. G. Kim, S. H. Bae, and H. Kim, "Carrier-phase-estimation algorithm featuring fast trackability for high-speed coherent WDM PON based on RSOA," *Optics Express*, vol. 25, pp. 14282–14289, June 2017.

[64] C. E. Shannon, "Communication in the Presence of Noise," *Proceedings of the IRE*, vol. 37, pp. 10–21, Jan. 1949.

[65] C. Henry, "Theory of the linewidth of semiconductor lasers," *IEEE Journal of Quantum Electronics*, vol. 18, pp. 259–264, Feb. 1982.

[66] Y. Yadin, M. Orenstein, and M. Shtaif, "Statistics of nonlinear phase noise in phase modulated fiber-optic communications systems," in *Optical Fiber Communication Conference, 2004. OFC 2004*, vol. 1, pp. 178–, Feb. 2004.

[67] Z. Tao, W. Yan, S. Oda, T. Hoshida, and J. C. Rasmussen, "A simplified model for nonlinear cross-phase modulation in hybrid optical coherent system," *Optics Express*, vol. 17, p. 13860, Aug. 2009.

[68] M. Secondini and E. Forestieri, "On XPM Mitigation in WDM Fiber-Optic Systems," *IEEE Photonics Technology Letters*, vol. 26, pp. 2252–2255, Nov. 2014.

[69] D. Lavery, M. Ionescu, S. Makovejs, E. Torrengo, and S. J. Savory, "A long-reach ultra-dense 10 Gbit/s WDM-PON using a digital coherent receiver," *Optics Express*, vol. 18, pp. 25855–25860, Dec. 2010.

[70] D. Lavery, R. Maher, D. S. Millar, B. C. Thomsen, P. Bayvel, and S. J. Savory, "Digital Coherent Receivers for Long-Reach Optical Access Networks," *Journal of Lightwave Technology*, vol. 31, pp. 609–620, Feb. 2013.

[71] R. Corsini, M. Presi, M. Artiglia, and E. Ciaramella, "10-Gb/s Long-Reach PON System With Low-Complexity Dispersion-Managed Coherent Receiver," *IEEE Photonics Journal*, vol. 7, pp. 1–8, Oct. 2015.

[72] D. Lavery, B. C. Thomsen, P. Bayvel, and S. J. Savory, "Reduced Complexity Equalization for Coherent Long-Reach Passive Optical Networks [Invited]," *Journal of Optical Communications and Networking*, vol. 7, p. A16, Jan. 2015.

[73] M. S. Erkılınç, D. Lavery, K. Shi, B. C. Thomsen, P. Bayvel, R. I. Killey, and S. J. Savory, "Polarization-Insensitive Single-Balanced Photodiode Coherent Receiver for Long-Reach WDM-PONs," *Journal of Lightwave Technology*, vol. 34, pp. 2034–2041, Apr. 2016.

[74] D. J. G. Mestdagh and C. M. Martin, "The super-PON concept and its technical challenges," in *Broadband Communications: Global infrastructure for the information age* (L. Mason and A. Casaca, eds.), IFIP Advances in Information and Communication Technology, pp. 333–345, Springer US, 1996.

[75] C. DeSanti, L. Du, C. Lam, and J. Jiang, "Super-pon: Scale Fully Passive Optical Access Networks to Longer Reaches and to a Significantly Higher Number of Subscribers." `http://www.ieee802.org/3/ad_hoc/ngrates/public/18_01/desanti_nea_01a_0118.pdf`, 2018. Accessed February 2020.

[76] A. Napoli, M. Bohn, D. Rafique, A. Stavdas, N. Sambo, L. Poti, M. Nölle, J. K. Fischer, E. Riccardi, A. Pagano, A. D. Giglio, M. S. Moreolo, J. M. Fabrega, E. Hugues-Salas, G. Zervas, D. Simeonidou, P. Layec, A. D'Errico, T. Rahman, and J. P. F. P. Giménez, "Next generation elastic optical networks: The vision of the European research project IDEALIST," *IEEE Communications Magazine*, vol. 53, pp. 152–162, Feb. 2015.

[77] J. E. Carroll, J. Whiteaway, and D. Plumb, *Distributed Feedback Semiconductor Lasers*. IET, Jan. 1998.

[78] K. Petermann, *Laser Diode Modulation and Noise*. Springer Science & Business Media, 7 1988.

[79] J. C. Cartledge and R. C. Srinivasan, "Extraction of DFB laser rate equation parameters for system simulation purposes," *Journal of Lightwave Technology*, vol. 15, pp. 852–860, May 1997.

[80] Q. Hu, D. Che, Y. Wang, F. Yuan, Q. Yang, and W. Shieh, "Complex modulation and detection with directly modulated lasers," *Opt. Express*, vol. 23, pp. 32809–32819, 12 2015.

[81] W. Franz, "Einfluß eines elektrischen feldes auf eine optische absorptionskante," *Zeitschrift für Naturforschung A*, vol. 13, no. 6, pp. 484–489, 2014.

[82] M. Chaciński, U. Westergren, L. Thylén, R. Schatz, J. Li, A. Djupsjöbacka, and B. Stoltz, "Modulation and chirp evaluation of 100 GHz DFB-TWEAM," in *36th European Conference and Exhibition on Optical Communication*, pp. 1–3, Sep. 2010.

[83] M. I. Olmedo, J. V. Olmos, U. Westergren, S. Popov, and I. T. Monroy, "Chirp investigation in EMLs towards frequency shift keying modulation," in *Asia Communications and Photonics Conference 2014*, p. ATh3A.12, Optical Society of America, 2014.

[84] M. Seimetz, *High-order modulation for optical fiber transmission*. No. 143 in Springer series in optical sciences, Springer, 2009. OCLC: ocn297148328.

[85] H. Yamazaki, S. Kanazawa, Y. Nakanishi, Y. Ueda, W. Kobayashi, Y. Muramoto, H. Ishii, and H. Sanjoh, "Ultra-broadband EA-DFB laser module for 200-Gbit/s PAM4 transmitter," in *2017 Optical Fiber Communications Conference and Exhibition (OFC)*, pp. 1–3, Mar. 2017.

[86] F. Li, Z. Li, Q. Sui, J. Li, X. Yi, L. Li, and Z. Li, "200 Gbit/s (68.25 Gbaud) PAM8 Signal Transmission and Reception for Intra-Data Center Interconnect," in *2019 Optical Fiber Communications Conference and Exhibition (OFC)*, pp. 1–3, Mar. 2019.

[87] M. Terayama, S. Okamoto, K. Kasai, M. Yoshida, and M. Nakazawa, "4096 QAM (72 Gbit/s) Single-Carrier Coherent Optical Transmission with a Potential SE of 15.8 bit/s/Hz in All-Raman Amplified 160 km Fiber Link," in *2018 Optical Fiber Communications Conference and Exposition (OFC)*, pp. 1–3, Mar. 2018.

[88] T. Miya, Y. Terunuma, T. Hosaka, and T. Miyashita, "Ultimate low-loss single-mode fibre at 1.55 μm," *Electronics Letters*, vol. 15, pp. 106–108, Feb. 1979.

[89] J. Gowar, *Optical Communication Systems (2nd Edition)*. Prentice Hall International - Series in Optoelectronics, 1993.

[90] G. P. Agrawal, *Fiber-optic communication systems*. New York: Wiley-Interscience, 3 ed., 2002. OCLC: 58437991.

[91] H. Yamazaki, M. Nagatani, H. Wakita, Y. Ogiso, M. Nakamura, M. Ida, H. Nosaka, T. Hashimoto, and Y. Miyamoto, "IMDD Transmission at Net Data Rate of 333 Gb/s Using Over-100-GHz-Bandwidth Analog Multiplexer and Mach–Zehnder Modulator," *Journal of Lightwave Technology*, vol. 37, pp. 1772–1778, Apr. 2019.

[92] K. Kikuchi, "Fundamentals of Coherent Optical Fiber Communications," *Journal of Lightwave Technology*, vol. 34, pp. 157–179, Jan. 2016.

[93] M. S. Faruk and S. J. Savory, "Digital Signal Processing for Coherent Transceivers Employing Multilevel Formats," *Journal of Lightwave Technology*, vol. 35, pp. 1125–1141, Mar. 2017.

[94] G. Agrawal, *Nonlinear Fiber Optics (3rd Edition)*. Academic press, 2001.

[95] S. J. Savory, G. Gavioli, R. I. Killey, and P. Bayvel, "Electronic compensation of chromatic dispersion using a digital coherent receiver," *Optics Express*, vol. 15, pp. 2120–2126, Mar. 2007.

[96] D. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems," *IEEE Transactions on Communications*, vol. 28, pp. 1867–1875, Nov. 1980.

[97] J. Yang, "The Multimodulus Blind Equalization and Its Generalized Algorithms," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 5, p. 19, 2002.

[98] I. Fatadin, D. Ives, and S. Savory, "Blind Equalization and Carrier Phase Recovery in a 16-QAM Optical Coherent System," *Journal of Lightwave Technology*, vol. 27, pp. 3042–3049, Aug. 2009.

[99] M. S. Faruk and K. Kikuchi, "Adaptive frequency-domain equalization in digital coherent optical receivers," *Optics Express*, vol. 19, pp. 12789–12798, June 2011.

[100] A. J. Viterbi and A. M. Viterbi, "Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission," *IEEE Transactions on Information Theory*, vol. 29, pp. 543–551, July 1983.

[101] E. Ip and J. M. Kahn, "Feedforward carrier recovery for coherent optical communications," *Journal of Lightwave Technology*, vol. 25, no. 9, pp. 2675–2692, 2007.

[102] M. G. Taylor, "Phase estimation methods for optical coherent detection using digital signal processing," *Journal of Lightwave Technology*, vol. 27, no. 7, pp. 901–914, 2009.

[103] G. Colavolpe and R. Raheli, "The capacity of noncoherent channels," in *1999 IEEE International Conference on Communications*, vol. 2, pp. 1182–1186 vol.2, June 1999.

[104] G. Colavolpe and R. Raheli, "Noncoherent sequence detection," *IEEE Transactions on Communications*, vol. 47, pp. 1376–1385, Sept. 1999.

[105] M. Katz and S. Shamai, "On the capacity-achieving distribution of the discrete-time non-coherent additive white Gaussian noise channel," in *Proceedings IEEE International Symposium on Information Theory,*, pp. 165–, June 2002.

[106] R. Nuriyev and A. Anastasopoulos, "Capacity characterization for the noncoherent block-independent AWGN channel," in *IEEE International Symposium on Information Theory, 2003. Proceedings.*, pp. 373–373, June 2003.

[107] S. O. Rice, "Statistical properties of a sine wave plus random noise," *The Bell System Technical Journal*, vol. 27, pp. 109–157, Jan. 1948.

[108] B. Goebel, R. Essiambre, G. Kramer, P. J. Winzer, and N. Hanik, "Calculation of Mutual Information for Partially Coherent Gaussian Channels With Applications to Fiber Optics," *IEEE Transactions on Information Theory*, vol. 57, pp. 5720–5736, Sept. 2011.

[109] P. Hou, B. J. Belzer, and T. R. Fischer, "Shaping gain of the partially coherent additive white Gaussian noise channel," *IEEE Communications Letters*, vol. 6, pp. 175–177, May 2002.

[110] K. V. Mardia and P. E. Jupp, *Directional statistics*. Wiley series in probability and statistics, Chichester ; New York: J. Wiley, 2000.

[111] "*NIST Digital Library of Mathematical Functions*." http://dlmf.nist.gov/, Release 1.0.22 of 2019-03-15. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds.

[112] G. J. Foschini, R. D. Gitlin, and S. B. Weinstein, "On the selection of a two-dimensional signal constellation in the presence of phase jitter and gaussian noise," *The Bell System Technical Journal*, vol. 52, no. 6, pp. 927–965, 1973.

[113] J. E. Volder, "The CORDIC Trigonometric Computing Technique," *IRE Transactions on Electronic Computers*, vol. EC-8, pp. 330–334, Sept. 1959.

[114] P. K. Meher, J. Valls, T. Juang, K. Sridharan, and K. Maharatna, "50 Years of CORDIC: Algorithms, Architectures, and Applications," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, pp. 1893–1907, Sept. 2009.

[115] S. D. N. Richaupadhyay, "Hardware Implementation Of Hyperbolic Tan Using Cordic On FPGA," *International Journal of Engineering Research and Applications (IJERA)*, vol. 3, pp. 696–699, Apr. 2014.

[116] B. Y. Masram and P. T. Karule, "High performance analysis of a CORDIC architectures based on FPGA: A comparative approach," in *2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies*, pp. 569–574, May 2014.

[117] A. Tolmachev, I. Tselniker, M. Meltsin, I. Sigron, D. Dahan, A. Shalom, and M. Nazarathy, "Multiplier-Free Phase Recovery With Polar-domain Multisymbol-Delay-Detector," *Journal of Lightwave Technology*, vol. 31, pp. 3638–3650, Dec. 2013.

[118] A. Alvarado, E. Agrell, D. Lavery, R. Maher, and P. Bayvel, "Replacing the soft-decision FEC limit paradigm in the design of optical communication systems," *Journal of Lightwave Technology*, vol. 34, no. 2, pp. 707–721, 2016.

[119] A. Alvarado, T. Fehenberger, B. Chen, and F. M. J. Willems, "Achievable information rates for fiber optics: Applications and computations," *Journal of Lightwave Technology*, vol. 36, no. 2, pp. 424–439, 2018.

[120] D. M. Arnold, H. Loeliger, P. O. Vontobel, A. Kavcic, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Transactions on Information Theory*, vol. 52, no. 8, pp. 3498–3508, 2006.

[121] S. Tretter, "Estimating the frequency of a noisy sinusoid by linear regression (Corresp.)," *IEEE Transactions on Information Theory*, vol. 31, pp. 832–835, Nov. 1985.

[122] H. Fu and P. Y. Kam, "MAP/ML Estimation of the Frequency and Phase of a Single Sinusoid in Noise," *IEEE Transactions on Signal Processing*, vol. 55, pp. 834–845, Mar. 2007.

[123] H. Fu and P. Y. Kam, "Exact phase noise model and its application to linear minimum variance estimation of frequency and phase of a noisy sinusoid," in *2008 IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–5, Sept. 2008.

[124] A. Leven, N. Kaneda, U. Koc, and Y. Chen, "Frequency estimation in intradyne reception," *IEEE Photonics Technology Letters*, vol. 19, no. 6, pp. 366–368, 2007.

[125] J. H. Ke, K. P. Zhong, Y. Gao, A. Bakhshali, and J. C. Cartledge, "Low-complexity joint frequency offset and carrier phase estimation based on QPSK partitioning for DP 16-QAM," in *2013 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC)*, pp. 1–3, Mar. 2013.

[126] R. Corvaja and S. Pupolin, "Phase noise effects in QAM systems," in *Proceedings of 8th International Symposium on Personal, Indoor and Mobile Radio Communications - PIMRC '97*, vol. 2, pp. 452–456 vol.2, Sept. 1997.

[127] T. Pfau, S. Hoffmann, and R. Noe, "Hardware-efficient coherent digital receiver concept with feedforward carrier recovery for m-QAM constellations," *Journal of Lightwave Technology*, vol. 27, no. 8, pp. 989–999, 2009.

[128] M. Sales Llopis, M. S. Faruk, and S. J. Savory, "Improved linewidth tolerant carrier phase recovery based on polar MAP metric estimate," in *Optical Fiber Communication Conference (2017), paper Th4C.2*, p. Th4C.2, Optical Society of America, 2017.

[129] W. Weber, "Differential encoding for multiple amplitude and phase shift keying systems," *IEEE Transactions on Communications*, vol. 26, no. 3, pp. 385–391, 1978.

[130] R. Noe, "PLL-free synchronous QPSK polarization multiplex/diversity receiver concept with digital IQ baseband processing," *IEEE Photonics Technology Letters*, vol. 17, no. 4, pp. 887–889, 2005.

[131] W. Webb and L. L. Hanzo, *Modern Quadrature Amplitude Modulation: Principles and Applications for Fixed and Wireless Communications*. I.E.E.E.Press, 1994.

[132] W. K. M. Ahmed and K. Balachandran, "Uncoded symbol error rate estimation: methods and analysis," *IEEE Transactions on Vehicular Technology*, vol. 54, pp. 1950–1962, Nov. 2005.

[133] Tianping Feng, Jian Chen, Zhaohui Ma, Yuanwei Fan, and Xiaodi You, "BER performance of bandwidth-limited QPSK optical coherent receiver with carrier phase estimation," in *2015 14th International Conference on Optical Communications and Networks (ICOCN)*, pp. 1–3, July 2015.

[134] H. Mahmoud and H. Arslan, "Error vector magnitude to SNR conversion for nondata-aided receivers," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 2694–2704, May 2009.

[135] M. Alvarez-Diaz, R. Lopez-Valcarce, and C. Mosquera, "SNR Estimation for Multilevel Constellations Using Higher-Order Moments," *IEEE Transactions on Signal Processing*, vol. 58, pp. 1515–1526, Mar. 2010.

[136] G. P. Agrawal, "Intensity dependence of the linewidth enhancement factor and its implications for semiconductor lasers," *IEEE Photonics Technology Letters*, vol. 1, pp. 212–214, Aug. 1989.

[137] Q. Hu, D. Che, Y. Wang, F. Yuan, Q. Yang, and W. Shieh, "Complex modulation and detection with directly modulated lasers," *Optics Express*, vol. 23, p. 32809, Dec. 2015.

[138] Q. Hu, D. Che, Y. Wang, A. Li, J. Fang, and W. Shieh, "Beyond amplitude-only detection for digital coherent system using directly modulated laser," *Optics Letters*, vol. 40, p. 2762, June 2015.

[139] D. Che, Q. Hu, F. Yuan, Q. Yang, and W. Shieh, "Enabling Complex Modulation of Directly Modulated Signals Using Laser Frequency Chirp," *IEEE Photonics Technology Letters*, vol. 27, pp. 2407–2410, Nov. 2015.

[140] D. Che, F. Yuan, Q. Hu, and W. Shieh, "Frequency Chirp Supported Complex Modulation of Directly Modulated Lasers," *Journal of Lightwave Technology*, vol. 34, pp. 1831–1836, Apr. 2016.

[141] G. Forney, "Maximum-likelihood sequence estimation of digital sequences in the presence of intersymbol interference," *IEEE Transactions on Information Theory*, vol. 18, pp. 363–378, May 1972.

[142] G. D. Forney, "The viterbi algorithm," *Proceedings of the IEEE*, vol. 61, pp. 268–278, Mar. 1973.

[143] A. Villafranca, J. Lasobras, and I. Garces, "Precise characterization of the frequency chirp in directly modulated DFB lasers," in *2007 Spanish Conference on Electron Devices*, pp. 173–176, Jan. 2007.

[144] G. P. Agrawal and M. J. Potasek, "Effect of frequency chirping on the performance of optical communication systems," *Optics Letters*, vol. 11, pp. 318–320, May 1986.

[145] M. N. Chughtai, *Nonlinear Phase Noise in Fiber Optical Communication*. PhD thesis, Chalmers University of Technology, 2009.

[146] J. P. Gordon and L. F. Mollenauer, "Phase noise in photonic communications systems using linear amplifiers," *Optics Letters*, vol. 15, pp. 1351–1353, Dec. 1990.

[147] Y. Yadin, M. Shtaif, and M. Orenstein, "Nonlinear phase noise in phase-modulated WDM fiber-optic communications," *IEEE Photonics Technology Letters*, vol. 16, pp. 1307–1309, May 2004.

[148] K.-P. Ho, "Probability density of nonlinear phase noise," *JOSA B*, vol. 20, pp. 1875–1879, Sept. 2003.

[149] A. P. T. Lau and J. M. Kahn, "Signal Design and Detection in Presence of Nonlinear Phase Noise," *Journal of Lightwave Technology*, vol. 25, pp. 3008–3016, Oct. 2007.

[150] T. Pfau, X. Liu, and S. Chandrasekhar, "Optimization of 16-ary quadrature amplitude modulation constellations for phase noise impaired channels," in *2011 37th European Conference and Exhibition on Optical Communication*, pp. 1–3, Sept. 2011. ISSN: 1550-381X.

[151] Z. H. Peric, I. B. Djordjevic, S. M. Bogosavljevic, and M. C. Stefanovic, "Design of signal constellations for Gaussian channel by using iterative polar quantization," in *MELECON '98. 9th Mediterranean Electrotechnical Conference. Proceedings (Cat. No.98CH36056)*, vol. 2, pp. 866–869, May 1998.

[152] M. P. Yankov, D. Zibar, K. J. Larsen, L. P. B. Christensen, and S. Forchhammer, "Constellation Shaping for Fiber-Optic Channels With QAM and High Spectral Efficiency," *IEEE Photonics Technology Letters*, vol. 26, pp. 2407–2410, Dec. 2014.

[153] R. T. Jones, T. A. Eriksson, M. P. Yankov, and D. Zibar, "Deep Learning of Geometric Constellation Shaping Including Fiber Nonlinearities," in *2018 European Conference on Optical Communication (ECOC)*, pp. 1–3, Sept. 2018.

[154] E. Agrell, A. Alvarado, G. Durisi, and M. Karlsson, "Capacity of a Nonlinear Optical Channel With Finite Memory," *Journal of Lightwave Technology*, vol. 32, pp. 2862–2876, Aug. 2014.

[155] D. Arnold and H. Loeliger, "On the information rate of binary-input channels with memory," in *ICC 2001. IEEE International Conference on Communications. Conference Record (Cat. No.01CH37240)*, vol. 9, pp. 2692–2695 vol.9, June 2001.

[156] P. Poggiolini, G. Bosco, and M. Visintin, "MLSE receivers and their applications in optical transmission systems," in *LEOS 2007 - IEEE Lasers and Electro-Optics Society Annual Meeting Conference Proceedings*, pp. 216–217, Oct. 2007. ISSN: 1092-8081.

[157] P. A. Morton, T. Tanbun-Ek, R. A. Logan, A. M. Sergent, P. F. Sciortino, and D. L. Coblentz, "Frequency response subtraction for simple measurement of intrinsic laser dynamic properties," *IEEE Photonics Technology Letters*, vol. 4, pp. 133–136, Feb. 1992.

[158] J. C. Cartledge and R. C. Srinivasan, "Extraction of DFB laser rate equation parameters for system simulation purposes," *Journal of Lightwave Technology*, vol. 15, pp. 852–860, May 1997.

[159] L. Bjerkan, A. Royset, L. Hafskjaer, and D. Myhre, "Measurement of laser parameters for simulation of high-speed fiberoptic systems," *Journal of Lightwave Technology*, vol. 14, pp. 839–850, May 1996.

[160] P. Andre, A. Teixeira, L. P. Pellegrino, M. Lima, N. R, P. Monteiro, A. N. Pinto, J. L. Pinto, and J. F. D. Rocha, "Extraction of laser parameters for simulation purposes," in *NUSOD '05. Proceedings of the 5th International Conference on Numerical Simulation of Optoelectronic Devices, 2005.*, pp. 103–104, Sept. 2005.

[161] I. Fatadin, D. Ives, and M. Wicks, "Numerical simulation of intensity and phase noise from extracted parameters for CW DFB lasers," *IEEE Journal of Quantum Electronics*, vol. 42, pp. 934–941, Sept. 2006.

[162] T. Schaer, R. Rusnov, S. Eagle, J. Jastrebski, S. Albanese, and X. Fernando, "A dynamic simulation model for semiconductor laser diodes," in *Electrical*

and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on, vol. 1, pp. 293–297, IEEE, 2003.

[163] B. L. Espe, *MATLAB Simulation of a Distributed Feedback (DFB) Laser with Chirp Effects*. PhD thesis, Naval Postgraduate School, 1994. Google-Books-ID: i32gcQAACAAJ.

[164] P. Corvini and T. Koch, "Computer simulation of high-bit-rate optical fiber transmission using single-frequency lasers," *Journal of Lightwave Technology*, vol. 5, pp. 1591–1595, Nov. 1987.

[165] F. Delpiano, R. Paoletti, P. Audagnotto, and M. Puleo, "High frequency modeling and characterization of high performance DFB laser modules," *IEEE Transactions on Components, Packaging, and Manufacturing Technology: Part B*, vol. 17, pp. 412–417, Aug. 1994.

[166] R. Tucker, "Large-signal circuit model for simulation of injection-laser modulation dynamics," *IEE Proceedings I - Solid-State and Electron Devices*, vol. 128, pp. 180–184, Oct. 1981.

[167] L. Illing and M. B. Kennel, "Shaping current waveforms for direct modulation of semiconductor lasers," *IEEE Journal of Quantum Electronics*, vol. 40, pp. 445–452, May 2004.

Appendix A

Stencil-aided SPSK constellation design

The static properties of the SPSK format and the equally spaced phase jumps, make it possible to use a graphical method to quickly find the power levels of the constellation, given the desired average power. To do so, it is required a plot in which the x- and y-axis represent power from symbols before and after a transition, respectively. Then, since for each pair of coordinates (x,y) there is a phase transition according to Equation (5.20), only the contour lines for the targeted values, $\pm\Delta\varphi_{SPSK,m}$ have to be drawn. These lines, converge asymptotically towards the origin (zero power), and diverge in outwards for increasing power. Figure A.1 shows two examples of such a plot, for a 4-SPSK, and for $\alpha = 5$ (left) and $\alpha = 8$ (right). It can be observed that the effect of α is that, for larger values, the contour lines become less divergent.

The phase contour lines are symmetrical with respect to the main diagonal labelled "0", being positive the jumps above it and negative below it. The 0-diagonal just means that if during a transition the power is unchanged, there will be no phase jump either. Additionally, a secondary x-axis is located on top of the square and is labelled "Average power". The values of this axis need to be scaled with respect to the bottom axis according to Equation (C.1) in Appendix C, that relates the average power with the lowest power point, P_0.

To use the resulting plot as a stencil, the procedure is as follows: initially choose the desired average signal power and locate it on the top x-axis. From there, trace down an imaginary vertical line and mark the point where it crosses the 0-

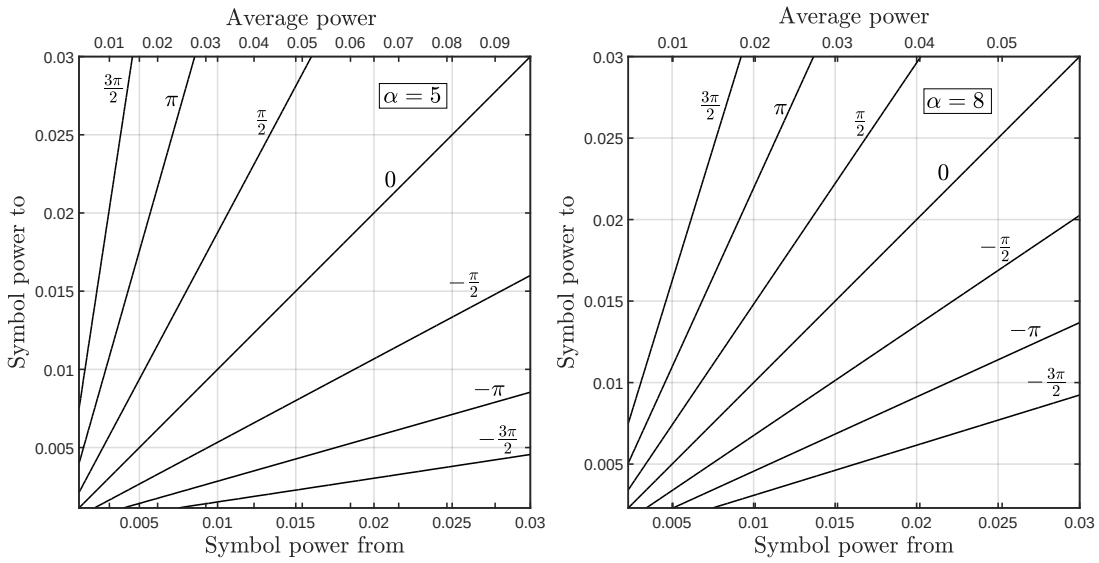


Figure A.1 – Construction of stencil for SPSK design. Example shown for $M = 4$ (4-SPSK) and two values of α: 5 (left) and 8 (right). Contour (straight) lines for all the phase transitions in the constellation. For higher values of α, the contour lines are closer to each other (smaller divergence)

diagonal: this is the starting point of the constellation, P_0. From this point, draw a vertical and a horizontal lines towards positive values (upwards and rightwards, respectively), until they reach the most positive and most negative contour lines. Then, close the squared defined by the drawn lines such that its bottom-left and top-right vertices lay on the 0-diagonal. Lastly, fill the square with a grid that connects all the crossing points at the contour lines, horizontally and vertically. Such grid will be composed of a total of M horizontal lines and M vertical lines, plus M^2 points, each representing the transition $P_{X\to Y}$. Figure A.2 illustrates an example of the stencil-based constellation design for a 4-SPSK with $\alpha = 5$. On the left, the light blue arrow connects the average power on top and the starting point P_0. From this point, the blue-shaded area emphasises the grid with 4 horizontal lines, 4 vertical lines, and 16 points. The 4 x-values of the vertical lines and the 4 y-values of the horizontal lines are identical, ans they represent the 4 power levels of the constellation. After this is done, and as shown in the right side of Figure A.2, it is easy to rescale the constellation graphically to take a different average power by just dragging the square grid along the 0-diagonal, upwards or downwards to increase or decrease the average power, respectively.

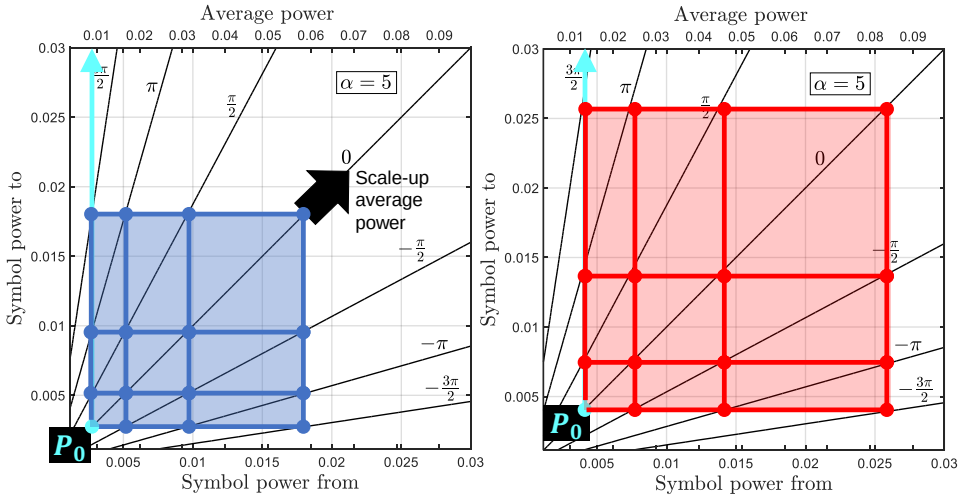


Figure A.2 – Example of use of the stencil for a 4-SPSK constellation design with a value of $\alpha = 5$.

Moreover, for small α, the contour lines become too divergent, making it difficult to appreciate the power crossing points. Plotting the stencil with logarithmic axes might be helpful, since the contour lines become parallel, as shown in Figure A.3 where an example of a 4-SPSK constellation is equivalently shown in both linear (left) and logarithmic (right) stencils, for a value of $\alpha = 4$.

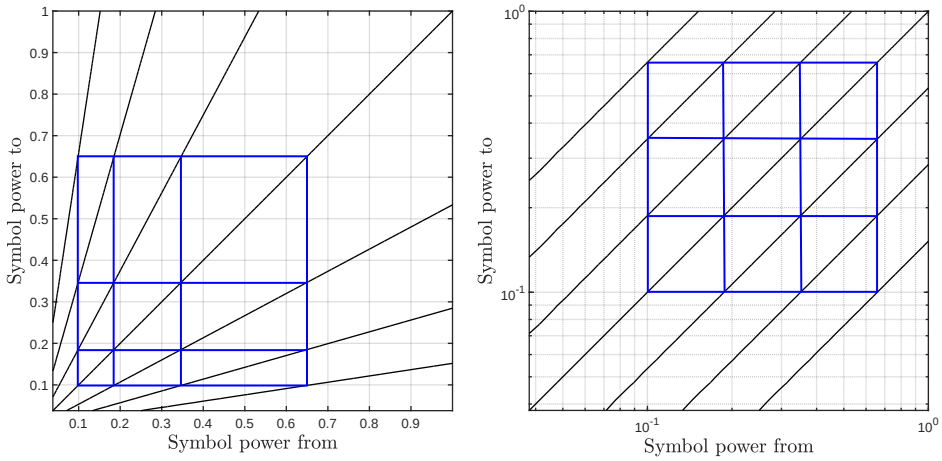


Figure A.3 – Equivalent 4-SPSK constellation design using a linear stencil (left) versus a logarithmic one (rigth).

Appendix B

On the max-min distance in 2-SAPM-D

Unlike with the conventional BPSK format, the maximum distance between two points on a spiral —limited to a 2π-radians loop— does not happen when the points are π radians apart. Instead, the angle between them that maximises the distance depends on the parameter α. Figure B.1 shows an example of three 2-SPAM-D constellations for different values of α: in (a), where $\alpha = 3$, the maximum distance happens where the points are at the same angle (0 for reference), but this is only so because the spiral has been limited up to this angle. In fact, the distance between the two points increases monotonically with the spiral and the choice of any pair of points would be according to a given laser and its power limitation. In (b), the point at maximum distance lies somewhere within the angles π and $3\pi/2$. As the value of α increases (like in (c) for $\alpha = 10$), the angle of maximum distance approaches π, converging to a BPSK form.

An interesting experiment consists of plotting the distance between a fixed point in the spiral, S_0 (the starting point), and a second point, S_x, as it moves away from S_0 along the spiral with the increase of the angle, $\Delta\varphi$. Such a plot is depicted in Figure B.2 where the distance versus the angle has been calculated for four different values of α. From these, two values of α are particularly remarkable: first, the curve for $\alpha \approx 3.66$ does not present a local maximum between angles π and 2π, but presents an inflexion point (see the local minimum tangential to the zero line on the plot below for ΔDistance). This means that for values of $\alpha < 3.66$ the distance between any two points in the spiral will be monotonically increasing with the angle

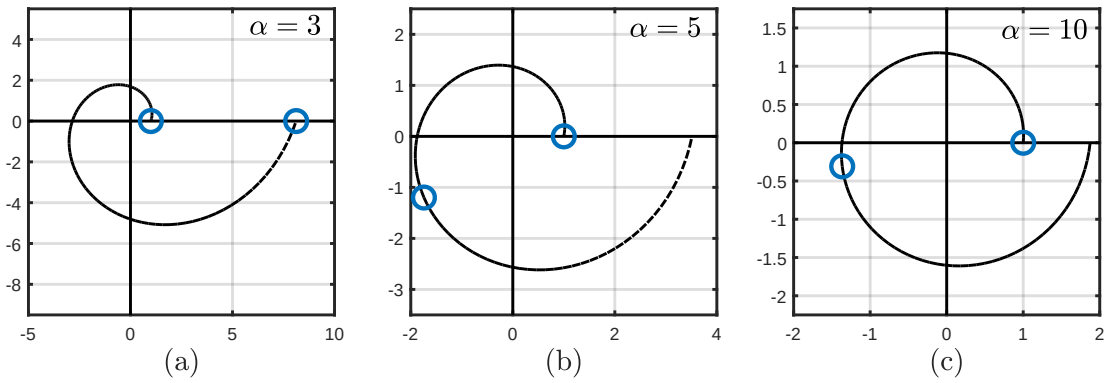


Figure B.1 – The maximum distance between two points on a spiral truncated at 2π radians depends on the parameter α. When α increases, the angle of maximum distance converges to π radians, as in a BPSK constellation.

and it would only make sense to choose a pair of constellation points if more data about the laser and its output power are given *a priori*. Secondly, the depicted curve for $\alpha = 4.33$ presents twice the same maximum distance at $\Delta\varphi \approx 4$ and $\Delta\varphi = 2\pi$ radians, respectively. This means that for α values above 3.66 and below 4.33, the spirals will present an absolute maximum at $\Delta\varphi = 2\pi$ and a smaller local maximum between 4 and 5 radians, and for $\alpha > 4.33$ there will be a single point with maximum distance located between π and 4 radians. Table B.1 summarises the possible mentioned cases.

Table B.1 – α boundaries in a 2-SPSK

Condition	$\Delta\varphi$ at max. distance (radians)
$\alpha \leq 3.66$	2π
$3.66 < \alpha \leq 4.33$	max at $\Delta\varphi = 2\pi$ and local max at $4 < \Delta\varphi < 5$
$\alpha > 4.33$	$\pi < \Delta\varphi < 4$

Therefore, the only calculation that might need to be carried out is to find the local maximum distance point for a given $\alpha > 3.66$ or the absolute maximum for $\alpha > 4.33$ when the spiral is limited to $\Delta\varphi \leq 2\pi$. Assuming a fixed starting point S_0 to have power $P_0 = 1$ and angle $\varphi = 0$, the distance between S_0 and any other point in the spiral S_x can be expressed as

$$D = |S_x - S_0| = \left|\sqrt{P_x}\exp\left(i\Delta\varphi\right) - \sqrt{1}\right| \tag{B.1}$$

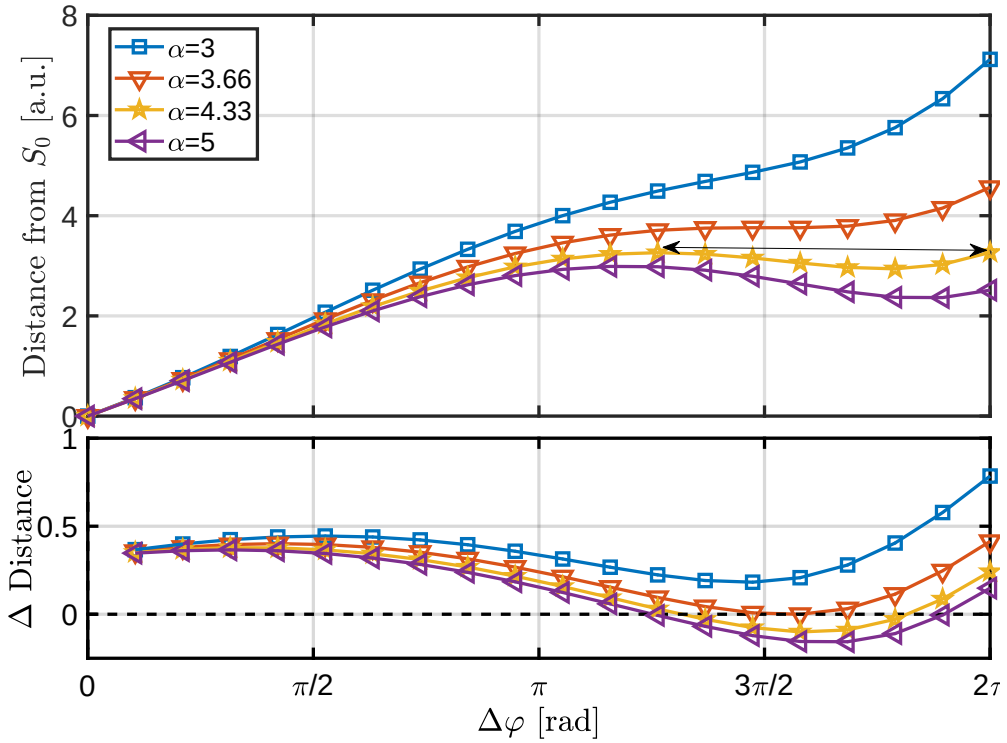


Figure B.2 – (top) Distance from a reference point S_0 and a rotated version S_x by $\Delta\varphi$ radians, for $\alpha = 3$ (monotonically increasing), 3.66 (first inflexion point), 4.33 (local maximum coincides with maximum at 2π radians, and 5 (only one maximum); (bottom) value of the slope of the Distance curve above (referred to as ΔDistance)

with $\Delta\varphi = \alpha/2 \ln(P_x)$. Although in the previous examples the spiral has been limited to 2π radians, in reality it will be limited by the maximum power that the laser output is restricted to. Thus, the local maximum for the distance has to be searched within the normalised range $(1, P_{max})$. However, maximising the distance in Equation (B.1) requires differentiating with respect to P_x and due to the imaginary argument of the exponential —showing an undulatory behaviour— it does not result in a closed-form solution. Instead, any of the multiple available numerical methods can be used to iteratively solve the problem. For this purpose, in this work the MATLAB function `fminsearch` has been used. Figure B.3 depicts the results obtained, showing, for different values of α, the ratio of the symbols power at maximum distance (top, for the absolute and local maximums within 2π radians) and the angle separation between symbols (bottom). It can be seen how, according to Table B.1, within the interval $3.66 < \alpha < 4.33$ the absolute and (secondary)

local maximum are not the same, but after $\alpha \geq 4.33$ these two become the same.

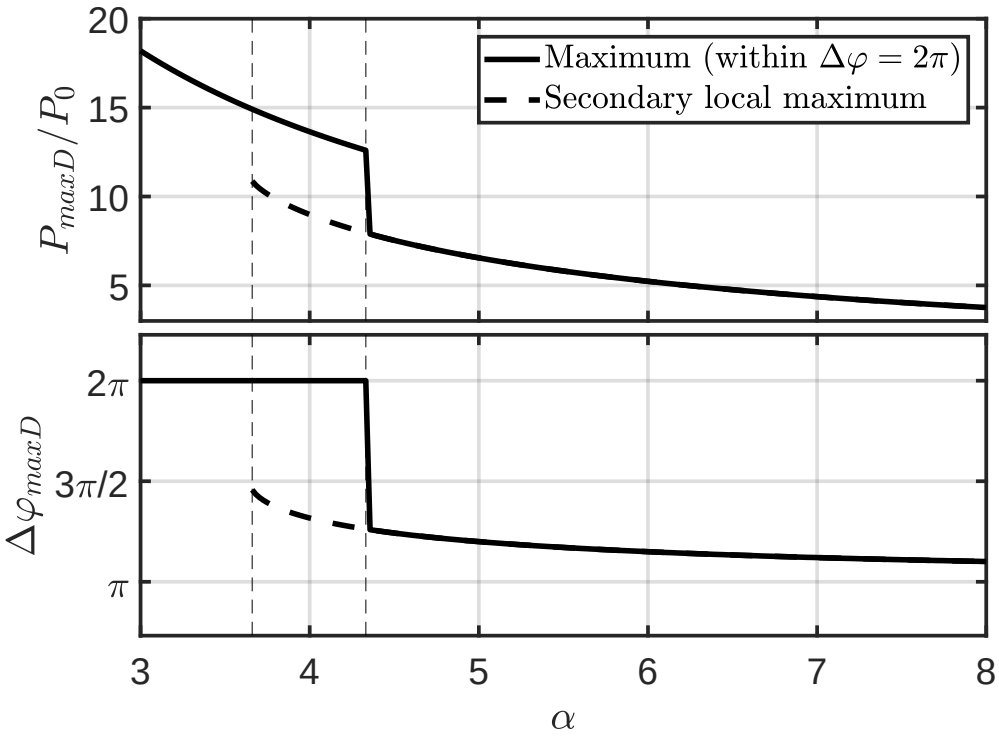


Figure B.3 – (top) Power ratio of the symbol that maximises the distance with respect to the reference point S_0 with $P_0 = 1$ and $\varphi_0 = 0$, P_{maxD}/P_0, versus the α parameter; (bottom) angle at which this maximisation occurs. A 2π-limited spiral has been considered.

Appendix C

Power implications of SPSK

C.1 Average power

It has been defined the parameter P_0 as the power of the "initial" point, or the point with the lowest power in the constellation. However, in practical applications it is the average signal power, $\langle P \rangle$, which is targeted to a specific value or within a range. So, in this case, for a given SPSK constellation set of parameters, and assuming that all the transmitted symbols are equally probable, the average power can be calculated as

$$\begin{aligned} \langle P_{SPSK} \rangle &= \frac{1}{M} \sum_{m=0}^{M-1} P_m \\ &= \frac{P_0}{M} \sum_{m=0}^{M-1} \exp\left(\frac{m\lambda 4\pi}{M\alpha}\right) \end{aligned} \tag{C.1}$$

and, therefore, after setting the average power to the desired value, $\langle P_{SPSK} \rangle = P_{avg}$, the reference value P_0 can be computed as

$$P_0 = M \cdot P_{avg} \left[\sum_{m=0}^{M-1} \exp\left(\frac{m\lambda 4\pi}{M\alpha}\right)\right]^{-1} \tag{C.2}$$

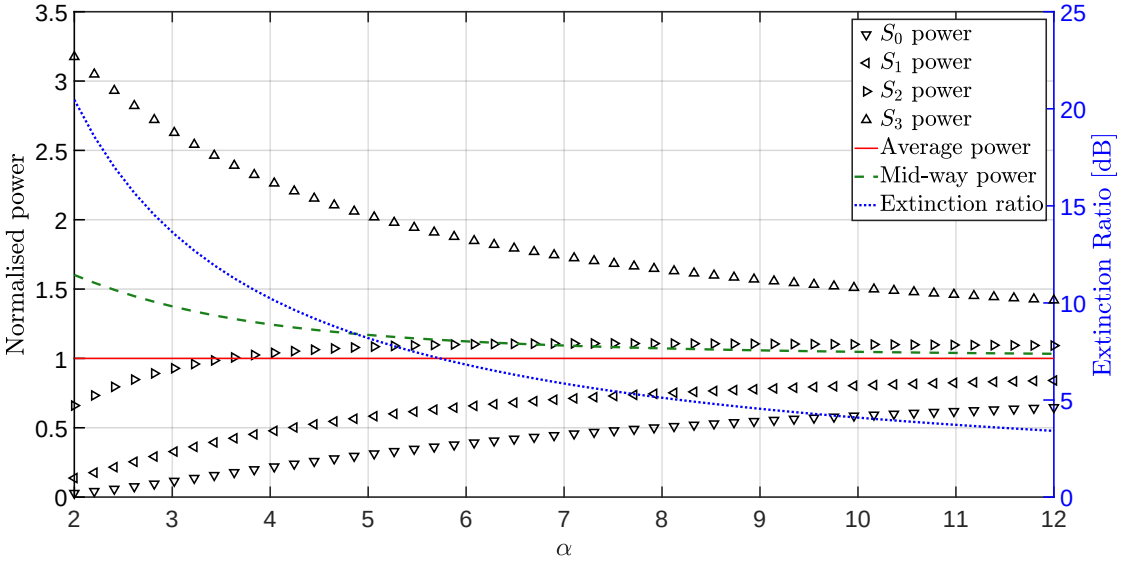


Figure C.1 – Power levels of a 4-SPSK constellation for symbols S_0 to S_3. The normalised average power is 1 (red solid line). The extinction ratio in dB (blue dotted line), decreases exponentially with α. The mid-way power (green dashed line) represents the reference at which the optical power modulation index (MI) would be symmetrical.

C.2 Extinction ratio

For the SPSK constellation, the extinction ratio (ER) can be calculated as the ratio between the powers of the last and first symbols, as

$$\begin{aligned} ER_{dB} &= 10\log_{10}\frac{P_{M-1}}{P_0} \\ &= 10\log_{10}\exp\left(\frac{(M-1)\,\lambda 4\pi}{M\alpha}\right) \\ &= \frac{10}{\ln 10}\frac{(M-1)\,\lambda 4\pi}{M\alpha} \end{aligned} \tag{C.3}$$

For a given SPSK constellation, the extinction ratio decreases exponentially with α. As an example Figure C.1 shows the extinction ratio in dB calculated for a 4-SPSK constellation (blue dotted line, referenced to the right y-axis).

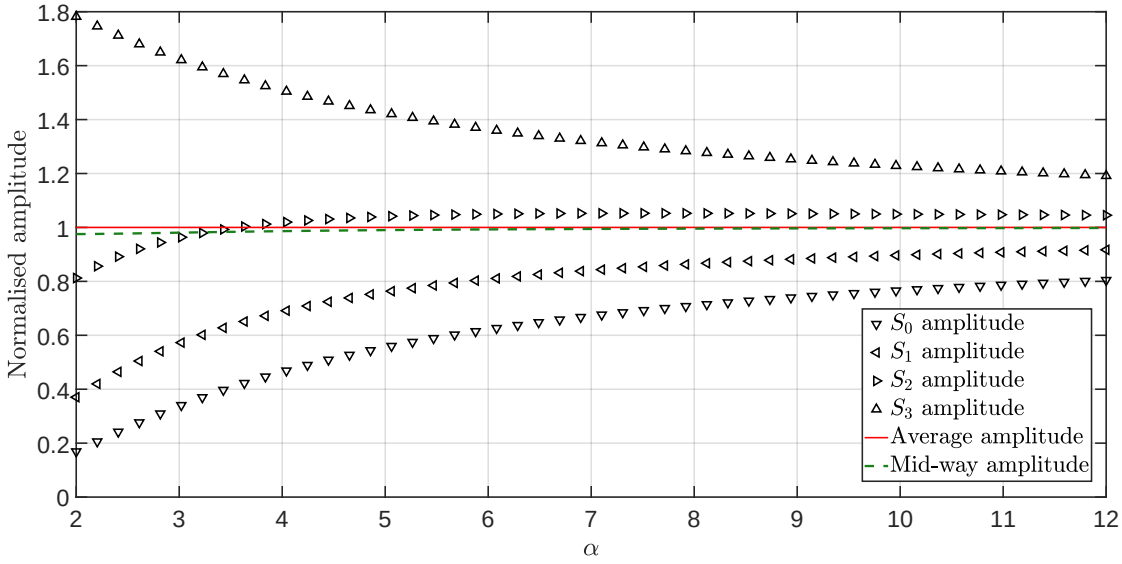


Figure C.2 – Amplitude levels of a 4-SPSK constellation (as the squared root of the power). The mid-way amplitude (green dashed line) is practically coincident with the average amplitude (red solid line) for any α, meaning that the modulation index (MI) can be considered symmetrical for the amplitude of electric field.

C.3 Modulation index

The modulation index is defined as the ratio of the maximum deviation of the modulated amplitude from the unmodulated carrier amplitude, usually expressed as a percentage. With laser sources, the modulation index shall be written in terms of modulated optical power, such that

$$MI = \frac{\max |P(t)|}{\langle P \rangle} \times 100 \tag{C.4}$$

In amplitude modulation (AM) systems, the index is symmetrical, i.e. it refers to deviations both above and under the unmodulated level. However, in the case of the SPSK modulation, this will not hold true due to the geometric progression that follow the consecutive power levels. This asymmetry is observed in Figure C.1 as how the normalised average power (red solid line) does not coincide with the mean value (mid-way power, dashed green line) between maximum and minimum symbol power levels. However, the mean power does asymptotically approach the average power as the parameter α increases. It is also noted from Figure C.2 that,

if electric field amplitude is used instead of power, for the same SPSK constellation, the mid-way point is practically equal to the average value, disregarding the value of α.